(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 September 2002 (19.09.2002)

PCT

(10) International Publication Number
WO 02/071928 A2

(51) International Patent Classification[7]: A61B

(21) International Application Number: PCT/US02/07826

(22) International Filing Date: 14 March 2002 (14.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/276,025	14 March 2001 (14.03.2001)	US
60/276,026	14 March 2001 (14.03.2001)	US
60/311,732	10 August 2001 (10.08.2001)	US
60/323,580	19 September 2001 (19.09.2001)	US
60/325,149	26 September 2001 (26.09.2001)	US
60/324,967	26 September 2001 (26.09.2001)	US
60/325,102	26 September 2001 (26.09.2001)	US

(71) Applicant *(for all designated States except US)*: **MILLENNIUM PHARMACEUTICALS, INC.** [US/US]; 75 Sidney Street, Cambridge, MA 02139 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **MONAHAN, John, E.** [US/US]; 942 West Street, Walpole, MA 02081 (US). **GANNAVARAPU, Manjula** [IN/US]; 10 Windemere Drive, Acton, MA 01720 (US). **HOERSCH, Sebastian** [US/US]; 127 Brattle Street, Arlington, MA 02474 (US). **KAMATKAR, Shubhangi** [US/US]; 655 Saw Mill Brook Parkway, Apt. 1, Newton, MA 02459 (US). **KOVATIS, Steven, G.** [US/US]; 94 Aldrich Road, Wilmington, MA 01887 (US). **MEYERS, Rachel, E.** [US/US]; 115 Devonshire Road, Newton, MA 02468 (US). **MORRISEY, Michael, P.** [US/US]; 140 Kenrick Street, Apt. 32, Brighton, MA 02135 (US). **OLANDT, Peter, J.** [US/US]; 29 Florence Street, Newton, MA 02459 (US). **SEN, Ami** [US/US]; 66 Dinsmore Avenue, Apt. 507, Framingham, MA 01702 (US). **VIEBY, Petter, Ole** [US/US]; 16 Nipmuck Drive, Westborough, MA 01581 (US). **MILLS, Gordon, B.** [CA/US]; 4124 Amherst Street, Houston, TX 77005 (US). **BAST, Robert, C., Jr.** [US/US]; 14 Memorial Point Lane, Houston, TX77024 (US). **LU, Karen** [US/US]; 4127 Amherst Street, Houston, TX 77005 (US). **SCHMANDT, Rosemarie, E.** [CA/US]; 7300 Brompton Road, Apt. 5512, Houston, TX 77025 (US). **ZHAO, Xumei** [US/US]; 6 Wildwood Lane, Burlington, MA 01803 (US). **GLATT, Karen** [US/US]; 17 Beacon Street, Natick, MA 01760 (US).

(74) Agents: **SMITH, DeAnn, F.** et al.; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/071928 A2

(54) Title: NUCLEIC ACID MOLECULES AND PROTEINS FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF OVARIAN CANCER

(57) Abstract: The invention relates to newly discovered nucleic acid molecules and proteins associated with ovarian cancer. Compositions, kits, and methods for detecting, characterizing, preventing, and treating human ovarian cancers are provided.

NUCLEIC ACID MOLECULES AND PROTEINS FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF OVARIAN CANCER

RELATED APPLICATIONS

The present application claims priority from U.S. provisional patent application serial no. 60/276,025, filed on March 14, 2001, which was abandoned on September 25, 2001, and from U.S. provisional patent application serial no. 60/325,149, filed on September 26, 2001. The present application also claims priority from U.S. provisional patent application serial no. 60/276,026, filed on March 14, 2001, which was abandoned on September 25, 2001, and from U.S. provisional patent application serial no. 60/324,967, filed September 26, 2001. The present application additionally claims priority from U.S. provisional patent application serial no. 60/311,732, filed August 10, 2001, which was abandoned on September 25, 2001, and from U.S. provisional patent application serial no. 60/325,102, filed September 26, 2001. The present application also claims priority from U.S. provisional patent application serial no. 60/323,580, filed September 19, 2001. All of the above applications are expressly incorporated by reference.

FIELD OF THE INVENTION

The field of the invention is ovarian cancer, including diagnosis, characterization, management, and therapy of ovarian cancer.

BACKGROUND OF THE INVENTION

Ovarian cancer is responsible for significant morbidity and mortality in populations around the world. Ovarian cancer is classified, on the basis of clinical and pathological features, in three groups, namely epithelial ovarian cancer (EOC; >90% of ovarian cancer in Western countries), germ cell tumors (*circa* 2-3% of ovarian cancer), and stromal ovarian cancer (*circa* 5% of ovarian cancer; Ozols *et al.*, 1997, *Cancer Principles and Practice of Oncology*, 5th ed., DeVita *et al.*, Eds. pp. 1502). Relative to EOC, germ cell tumors and stromal ovarian cancers are more easily detected and treated

at an early stage, translating into higher/better survival rates for patients afflicted with these two types of ovarian cancer.

There are numerous types of ovarian tumors, some of which are benign, and others of which are malignant. Treatment (including non-treatment) options and predictions of patient outcome depend on accurate classification of the ovarian cancer. Ovarian cancers are named according to the type of cells from which the cancer is derived and whether the ovarian cancer is benign or malignant. Recognized histological tumor types include, for example, serous, mucinous, endometrioid, and clear cell tumors. In addition, ovarian cancers are classified according to recognized grade and stage scales.

In grade I, the tumor tissue is well differentiated. In grade II, tumor tissue is moderately well differentiated. In grade III, the tumor tissue is poorly differentiated. This grade correlates with a less favorable prognosis than grades I and II. Stage I is generally confined within the capsule surrounding one (stage IA) or both (stage IB) ovaries, although in some stage I (*i.e.* stage IC) cancers, malignant cells may be detected in ascites, in peritoneal rinse fluid, or on the surface of the ovaries. Stage II involves extension or metastasis of the tumor from one or both ovaries to other pelvic structures. In stage IIA, the tumor extends or has metastasized to the uterus, the fallopian tubes, or both. Stage IIB involves extension of the tumor to the pelvis. Stage IIC is stage IIA or IIB in which malignant cells may be detected in ascites, in peritoneal rinse fluid, or on the surface of the ovaries. In stage III, the tumor comprises at least one malignant extension to the small bowel or the omentum, has formed extrapelvic peritoneal implants of microscopic (stage IIIA) or macroscopic (< 2 centimeter diameter, stage IIIB; > 2 centimeter diameter, stage IIIC) size, or has metastasized to a retroperitoneal or inguinal lymph node (an alternate indicator of stage IIIC). In stage IV, distant (*i.e.* non-peritoneal) metastases of the tumor can be detected.

The durations of the various stages of ovarian cancer are not presently known, but are believed to be at least about a year each (Richart *et al.*, 1969, *Am. J. Obstet. Gynecol.* 105:386). Prognosis declines with increasing stage designation. For example, 5-year survival rates for patients diagnosed with stage I, II, III, and IV ovarian cancer are 80%, 57%, 25%, and 8%, respectively.

Despite being the third most prevalent gynecological cancer, ovarian cancer is the leading cause of death among those afflicted with gynecological cancers. The disproportionate mortality of ovarian cancer is attributable to a substantial absence of symptoms among those afflicted with early-stage ovarian cancer and to difficulty diagnosing ovarian cancer at an early stage. Patients afflicted with ovarian cancer most often present with non-specific complaints, such as abnormal vaginal bleeding, gastrointestinal symptoms, urinary tract symptoms, lower abdominal pain, and generalized abdominal distension. These patients rarely present with paraneoplastic symptoms or with symptoms which clearly indicate their affliction. Presently, less than about 40% of patients afflicted with ovarian cancer present with stage I or stage II. Management of ovarian cancer would be significantly enhanced if the disease could be detected at an earlier stage, when treatments are much more generally efficacious.

Ovarian cancer may be diagnosed, in part, by collecting a routine medical history from a patient and by performing physical examination, x-ray examination, and chemical and hematological studies on the patient. Hematological tests which may be indicative of ovarian cancer in a patient include analyses of serum levels of proteins designated CA125 and DF3 and plasma levels of lysophosphatidic acid (LPA). Palpation of the ovaries and ultrasound techniques (particularly including endovaginal ultrasound and color Doppler flow ultrasound techniques) can aid detection of ovarian tumors and differentiation of ovarian cancer from benign ovarian cysts. However, a definitive diagnosis of ovarian cancer typically requires performing exploratory laparotomy of the patient.

Potential tests for the detection of ovarian cancer (*e.g.*, screening, reflex or monitoring) may be characterized by a number of factors. The "sensitivity" of an assay refers to the probability that the test will yield a positive result in an individual afflicted with ovarian cancer. The "specificity" of an assay refers to the probability that the test will yield a negative result in an individual not afflicted with ovarian cancer. The "positive predictive value" (PPV) of an assay is the ratio of true positive results (*i.e.* positive assay results for patients afflicted with ovarian cancer) to all positive results (*i.e.* positive assay results for patients afflicted with ovarian cancer + positive assay results for patients not afflicted with ovarian cancer). It has been estimated that in order for an assay to be an appropriate population-wide screening tool for ovarian cancer the

assay must have a PPV of at least about 10% (Rosenthal *et al.*, 1998, *Sem. Oncol.* 25:315-325). It would thus be desirable for a screening assay for detecting ovarian cancer in patients to have a high sensitivity and a high PPV. Monitoring and reflex tests would also require appropriate specifications.

Owing to the cost, limited sensitivity, and limited specificity of known methods of detecting ovarian cancer, screening is not presently performed for the general population. In addition, the need to perform laparotomy in order to diagnose ovarian cancer in patients who screen positive for indications of ovarian cancer limits the desirability of population-wide screening, such that a PPV even greater than 10% would be desirable.

Prior use of serum CA125 level as a diagnostic marker for ovarian cancer indicated that this method exhibited insufficient specificity for use as a general screening method. Use of a refined algorithm for interpreting CA125 levels in serial retrospective samples obtained from patients improved the specificity of the method without shifting detection of ovarian cancer to an earlier stage (Skakes, 1995, *Cancer* 76:2004). Screening for LPA to detect gynecological cancers including ovarian cancer exhibited a sensitivity of about 96% and a specificity of about 89%. However, CA125-based screening methods and LPA-based screening methods are hampered by the presence of CA125 and LPA, respectively, in the serum of patients afflicted with conditions other than ovarian cancer. For example, serum CA125 levels are known to be associated with menstruation, pregnancy, gastrointestinal and hepatic conditions such as colitis and cirrhosis, pericarditis, renal disease, and various non-ovarian malignancies. Serum LPA is known, for example, to be affected by the presence of non-ovarian gynecological malignancies. A screening method having a greater specificity for ovarian cancer than the current screening methods for CA125 and LPA could provide a population-wide screening for early stage ovarian cancer.

Presently greater than about 60% of ovarian cancers diagnosed in patients are stage III or stage IV cancers. Treatment at these stages is largely limited to cytoreductive surgery (when feasible) and chemotherapy, both of which aim to slow the spread and development of metastasized tumor. Substantially all late stage ovarian cancer patients currently undergo combination chemotherapy as primary treatment, usually a combination of a platinum compound and a taxane. Median survival for

responding patients is about one year. Combination chemotherapy involving agents such as doxorubicin, cyclophosphamide, cisplatin, hexamethylmelamine, paclitaxel, and methotrexate may improve survival rates in these groups, relative to single-agent therapies. Various recently-developed chemotherapeutic agents and treatment regimens have also demonstrated usefulness for treatment of advanced ovarian cancer. For example, use of the topoisomerase I inhibitor topectan, use of amifostine to minimize chemotherapeutic side effects, and use of intraperitoneal chemotherapy for patients having peritoneally implanted tumors have demonstrated at least limited utility. Presently, however, the 5-year survival rate for patients afflicted with stage III ovarian cancer is 25%, and the survival rate for patients afflicted with stage IV ovarian cancer is 8%.

In summary, the earlier ovarian cancer is detected, the aggressiveness of therapeutic intervention and the side effects associated with therapeutic intervention are minimized. More importantly, the earlier the cancer is detected, the survival rate and quality of life of ovarian cancer patients is enhanced. Thus, a pressing need exists for methods of detecting ovarian cancer as early as possible. There also exists a need for methods of detecting recurrence of ovarian cancer as well as methods for predicting and monitoring the efficacy of treatment. There further exists a need for new therapeutic methods for treating ovarian cancer. The present invention satisfies these needs.

SUMMARY OF THE INVENTION

The invention relates to cancer markers (hereinafter "markers" or "markers of the inventions"), which are listed in Tables 1-3. The invention provides nucleic acids and proteins that are encoded by or correspond to the markers (hereinafter "marker nucleic acids" and "marker proteins," respectively). The invention further provides antibodies, antibody derivatives and antibody fragments which bind specifically with such proteins and/or fragments of the proteins.

In one aspect, the invention relates to various diagnostic, monitoring, test and other methods related to ovarian cancer detection and therapy. In one embodiment, the invention provides a diagnostic method of assessing whether a patient has ovarian cancer or has higher than normal risk for developing ovarian cancer, comprising the steps of comparing the level of expression of a marker of the invention in a patient

sample and the normal level of expression of the marker in a control, *e.g.*, a sample from a patient without ovarian cancer. A significantly higher level of expression of the marker in the patient sample as compared to the normal level is an indication that the patient is afflicted with ovarian cancer or has higher than normal risk for developing ovarian cancer.

In a preferred embodiment of the diagnostic method, the marker is over-expressed by at least two-fold in at least about 20% of stage I ovarian cancer patients, stage II ovarian cancer patients, stage III ovarian cancer patients, stage IV ovarian cancer patients, grade I ovarian cancer patients, grade II ovarian cancer patients, grade III ovarian cancer patients, epithelial ovarian cancer patients, stromal ovarian cancer patients, germ cell ovarian cancer patients, malignant ovarian cancer patients, benign ovarian cancer patients, serous neoplasm ovarian cancer patients, mucinous neoplasm ovarian cancer patients, endometrioid neoplasm ovarian cancer patients and/or clear cell neoplasm ovarian cancer patients.

The diagnostic methods of the present invention are particularly useful for patients with an identified pelvic mass or symptoms associated with ovarian cancer. The methods of the present invention can also be of particular use with patients having an enhanced risk of developing ovarian cancer (*e.g.*, patients having a familial history of ovarian cancer, patients identified as having a mutant oncogene, and patients at least about 50 years of age).

In a preferred diagnostic method of assessing whether a patient is afflicted with ovarian cancer (*e.g.*, new detection ("screening"), detection of recurrence, reflex testing), the method comprises comparing:

a) the level of expression of a marker of the invention in a patient sample, and

b) the normal level of expression of the marker in a control non-ovarian cancer sample.

A significantly higher level of expression of the marker in the patient sample as compared to the normal level is an indication that the patient is afflicted with ovarian cancer.

The invention also provides diagnostic methods for assessing the efficacy of a therapy for inhibiting ovarian cancer in a patient. Such methods comprise comparing:

a) expression of a marker of the invention in a first sample obtained from the patient prior to providing at least a portion of the therapy to the patient, and

b) expression of the marker in a second sample obtained from the patient following provision of the portion of the therapy.

A significantly lower level of expression of the marker in the second sample relative to that in the first sample is an indication that the therapy is efficacious for inhibiting ovarian cancer in the patient.

It will be appreciated that in these methods the "therapy" may be any therapy for treating ovarian cancer including, but not limited to, chemotherapy, radiation therapy, surgical removal of tumor tissue, gene therapy and biologic therapy such as the administering of antibodies and chemokines. Thus, the methods of the invention may be used to evaluate a patient before, during and after therapy, for example, to evaluate the reduction in tumor burden.

In a preferred embodiment, the diagnostic methods of the present invention are directed to therapy using a chemical or biologic agent. These methods comprise comparing:

a) expression of a marker of the invention in a first sample obtained from the patient and maintained in the presence of the chemical or biologic agent, and

b) expression of the marker in a second sample obtained from the patient and maintained in the absence of the agent.

A significantly lower level of expression of the marker in the first sample relative to that in the second sample is an indication that the agent is efficacious for inhibiting ovarian cancer in the patient. In one embodiment, the first and second samples can be portions of a single sample obtained from the patient or portions of pooled samples obtained from the patient.

The invention additionally provides a monitoring method for assessing the progression of ovarian cancer in a patient, the method comprising:

a) detecting in a patient sample at a first time point, the expression of a marker of the invention;

b) repeating step a) at a subsequent time point in time; and

c) comparing the level of expression detected in steps a) and b), and therefrom monitoring the progression of ovarian cancer in the patient.

A significantly higher level of expression of the marker in the sample at the subsequent time point from that of the sample at the first time point is an indication that the ovarian cancer has progressed, whereas a significantly lower level of expression is an indication that the ovarian cancer has regressed.

The invention further provides a diagnostic method for determining whether ovarian cancer has metastasized or is likely to metastasize in the future, the method comprising comparing:

a) the level of expression of a marker of the invention in a patient sample, and

b) the normal level (or non-metastatic level) of expression of the marker in a control sample.

A significantly higher level of expression in the patient sample as compared to the normal level (or non-metastatic level) is an indication that the ovarian cancer has metastasized or is likely to metastasize in the future.

The invention moreover provides a test method for selecting a composition for inhibiting ovarian cancer in a patient. This method comprises the steps of:

a) obtaining a sample comprising cancer cells from the patient;

b) separately maintaining aliquots of the sample in the presence of a plurality of test compositions;

c) comparing expression of a marker of the invention in each of the aliquots; and

d) selecting one of the test compositions which significantly reduces the level of expression of the marker in the aliquot containing that test

composition, relative to the levels of expression of the marker in the presence of the other test compositions.

The invention additionally provides a test method of assessing the ovarian carcinogenic potential of a compound. This method comprises the steps of:

a) maintaining separate aliquots of ovarian cells in the presence and absence of the compound; and

b) comparing expression of a marker of the invention in each of the aliquots.

A significantly higher level of expression of the marker in the aliquot maintained in the presence of the compound, relative to that of the aliquot maintained in the absence of the compound, is an indication that the compound possesses ovarian carcinogenic potential.

In addition, the invention further provides a method of inhibiting ovarian cancer in a patient. This method comprises the steps of:

a) obtaining a sample comprising cancer cells from the patient;

b) separately maintaining aliquots of the sample in the presence of a plurality of compositions;

c) comparing expression of a marker of the invention in each of the aliquots; and

d) administering to the patient at least one of the compositions which significantly lowers the level of expression of the marker in the aliquot containing that composition, relative to the levels of expression of the marker in the presence of the other compositions.

In the aforementioned methods, the samples or patient samples comprise cells obtained from the patient. The cells may be found in an ovarian tissue sample collected, for example, by an ovarian tissue biopsy or histology section. In one embodiment, the patient sample is an ovary-associated body fluid. Such fluids include, for example, blood fluids, lymph, ascites fluids, gynecological fluids, cystic fluids, urine, and fluids collected by peritoneal rinsing. In another embodiment, the sample comprises cells obtained from the patient. In this embodiment, the cells may be found in a fluid selected from the group consisting of a fluid collected by peritoneal rinsing, a fluid collected by uterine rinsing, a uterine fluid, a uterine exudate, a pleural fluid, and an ovarian exudate. In a further embodiment, the patient sample is *in vivo*.

According to the invention, the level of expression of a marker of the invention in a sample can be assessed, for example, by detecting the presence in the sample of:

- the corresponding marker protein or a fragment of the protein (*e.g.* by using a reagent, such as an antibody, an antibody derivative, an antibody fragment or single-chain antibody, which binds specifically with the protein or protein fragment).
- the corresponding marker nucleic acid or a fragment of the nucleic acid (*e.g.* by contacting transcribed polynucleotides obtained from the sample with a substrate having affixed thereto one or more nucleic acids having the entire or a segment of the sequence or a complement thereof)
- a metabolite which is produced directly (*i.e.*, catalyzed) or indirectly by the corresponding marker protein.

According to the invention, any of the aforementioned methods may be performed using a plurality (*e.g.* 2, 3, 5, or 10 or more) of ovarian cancer markers, including ovarian cancer markers known in the art. In such methods, the level of expression in the sample of each of a plurality of markers, at least one of which is a marker of the invention, is compared with the normal level of expression of each of the plurality of markers in samples of the same type obtained from control humans not afflicted with ovarian cancer. A significantly altered (*i.e.*, increased or decreased as specified in the above-described methods using a single marker) level of expression in the sample of one or more markers of the invention, or some combination thereof, relative to that marker's corresponding normal levels, is an indication that the patient is afflicted with ovarian cancer. For all of the aforementioned methods, the marker(s) are preferably selected such that the positive predictive value of the method is at least about 10%.

In a further aspect, the invention provides an antibody, an antibody derivative, or an antibody fragment, which binds specifically with a marker protein or a fragment of the protein. The invention also provides methods for making such antibody, antibody derivative, and antibody fragment. Such methods may comprise immunizing a mammal with a protein or peptide comprising the entirety, or a segment of 10 amino acids or more, of a marker protein, wherein the protein or peptide may be obtained from

a cell or by chemical synthesis. The methods of the invention also encompass producing monoclonal and single-chain antibodies, which would further comprise isolating splenocytes from the immunized mammal, fusing the isolated splenocytes with an immortalized cell line to form hybridomas, and screening individual hybridomas for those that produce an antibody that binds specifically with a marker protein or a fragment of the protein.

In another aspect, the invention relates to various diagnostic and test kits. In one embodiment, the invention provides a kit for assessing whether a patient is afflicted with ovarian cancer. The kit comprises a reagent for assessing expression of a marker of the invention. In another embodiment, the invention provides a kit for assessing the suitability of a chemical or biologic agent for inhibiting an ovarian cancer in a patient. Such kit comprises a reagent for assessing expression of a marker of the invention, and may also comprise one or more of such agents. In a further embodiment, the invention provides kits for assessing the presence of ovarian cancer cells or treating ovarian cancers. Such kits comprise an antibody, an antibody derivative, or an antibody fragment, which binds specifically with a marker protein, or a fragment of the protein. Such kits may also comprise a plurality of antibodies, antibody derivatives, or antibody fragments wherein the plurality of such antibody agents binds specifically with a marker protein, or a fragment of the protein.

In an additional embodiment, the invention also provides a kit for assessing the presence of ovarian cancer cells, wherein the kit comprises a nucleic acid probe that binds specifically with a marker nucleic acid or a fragment of the nucleic acid. The kit may also comprise a plurality of probes, wherein each of the probes binds specifically with a marker nucleic acid, or a fragment of the nucleic acid.

In a further aspect, the invention relates to methods for treating a patient afflicted with ovarian cancer or at risk of developing ovarian cancer. Such methods may comprise reducing the expression and/or interfering with the biological function of a marker of the invention. In one embodiment, the method comprises providing to the patient an antisense oligonucleotide or polynucleotide complementary to a marker nucleic acid, or a segment thereof. For example, an antisense polynucleotide may be provided to the patient through the delivery of a vector that expresses an antisense polynucleotide of a marker nucleic acid or a fragment thereof. In another embodiment,

the method comprises providing to the patient an antibody, an antibody derivative, or antibody fragment, which binds specifically with a marker protein or a fragment of the protein. In a preferred embodiment, the antibody, antibody derivative or antibody fragment binds specifically with a protein having the sequence of any of the markers listed in Table 1, or a fragment of such a protein.

It will be appreciated that the methods and kits of the present invention may also include known cancer markers including known ovarian cancer markers. It will further be appreciated that the methods and kits may be used to identify cancers other than ovarian cancer.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts a graph which represents the results of the TaqMan® expression study.

DETAILED DESCRIPTION OF THE INVENTION

The invention relates to newly discovered markers, identified in Tables 1-3, that are associated with the cancerous state of ovarian cells. It has been discovered that the higher than normal level of expression of any of these markers or combination of these markers correlates with the presence of ovarian cancer in a patient. Methods are provided for detecting the presence of ovarian cancer in a sample, the absence of ovarian cancer in a sample, the stage of an ovarian cancer, and with other characteristics of ovarian cancer that are relevant to prevention, diagnosis, characterization, and therapy of ovarian cancer in a patient. Methods of treating ovarian cancer are also provided.

Tables 1-3 list the markers of the present invention. In the Tables the markers are identified with a name ("Marker"), the name the gene is commonly known by, if applicable ("Gene Name"), the Sequence Listing identifier of the cDNA sequence of a nucleotide transcript encoded by or corresponding to the marker ("SEQ ID NO (nts)"), the Sequence Listing identifier of the amino acid sequence of a protein encoded by the nucleotide transcript ("SEQ ID NO (AAs)"), and the location of the protein coding sequence within the cDNA sequence ("CDS").

Table 1 lists all of the markers of the invention, which are over-expressed in ovarian cancer cells compared to normal (*i.e.*, non-cancerous) ovarian cells and comprises markers listed in Tables 2 and 3. Table 2 lists newly-identified nucleotide

and amino acid sequences useful as ovarian cancer markers. Table 3 lists newly-identified nucleotide sequences useful as ovarian cancer markers.

In addition to their use in ovarian cancer, it has been found that the markers of the present invention may be used in the diagnosis, characterization, management, and therapy of additional diseases. For example, OV65 (SEQ ID NOS: 305 and 306), M593 (SEQ ID NOS: 307 and 308) and M594 (SEQ ID NOS: 309 and 310), are spondin molecules, and have one or more of the following activities: (1) neural cell adhesion and (2) neurite extension and can thus be used in, for example, the diagnosis and treatment of brain and CNS related disorders. Such brain and CNS related disorders include, but are not limited to, bacterial and viral meningitis, Alzheimers Disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers (*e.g.*, metastatic carcinoma of the brain, glioblastoma, lymphoma, astrocytoma, acoustic neuroma), hydrocephalus, and encephalitis. In another example, OV65, M593 and M594 polypeptides, nucleic acids, and modulators thereof can be used to treat disorders of the brain, such as cerebral edema, hydrocephalus, brain herniations, iatrogenic disease (due to, *e.g.*, infection, toxins, or drugs), inflammations (*e.g.*, bacterial and viral meningitis, encephalitis, and cerebral toxoplasmosis), cerebrovascular diseases (*e.g.*, hypoxia, ischemia, infarction, intracranial hemorrhage, vascular malformations, and hypertensive encephalopathy), and tumors (*e.g.*, neuroglial tumors, neuronal tumors, tumors of pineal cells, meningeal tumors, primary and secondary lymphomas, intracranial tumors, and medulloblastoma), and to treat injury or trauma to the brain.

OV25 (SEQ ID NOS: 360 and 361), an HE4 protein, has one or more of the following activities: (1) sperm maturation and (2) inhibition of extracellular proteases and can thus be used in, for example, the treatment and diagnosis of diseases and disorders relating to spermatogenesis. For example, OV25 polypeptides, nucleic acids, and modulators thereof can be used to treat testicular disorders, such as unilateral testicular enlargement (*e.g.*, nontuberculous, granulomatous orchitis); inflammatory diseases resulting in testicular dysfunction (*e.g.*, gonorrhea and mumps); cryptorchidism; sperm cell disorders (*e.g.*, immotile cilia syndrome and germinal cell aplasia); acquired testicular defects (*e.g.*, viral orchitis); and tumors (*e.g.*, germ cell tumors, interstitial cell tumors, androblastoma, testicular lymphoma and adenomatoid tumors).

OV52 (SEQ ID NOS: 190 and 191), a Pump-1 proteinase, has been found to have one or more of the following activities: (1) breakdown of extracellular matrix in normal physiological processes, such as embryonic development, reproduction, and remodeling, as well as in (2) disease processes, such as arthritis, and metastasis. Hence, OV52 nucleic acids, proteins, and modulators thereof can be used to modulate disorders associated with adhesion and migration of cells, *e.g.*, platelet aggregation disorders (*e.g.*, Glanzmann's thromboasthemia, which is a bleeding disorder characterized by failure of platelet aggregation in response to cell stimuli), inflammatory disorders (*e.g.*, leukocyte adhesion deficiency, which is a disorder associated with impaired migration of neutrophils to sites of extravascular inflammation), connective tissue disorders, arthritis, disorders associated with abnormal tissue migration during embryo development, and tumor metastasis.

M604 (SEQ ID NOS: 48 and 49), OV10 (SEQ ID NOS: 50 and 51), and M360 (SEQ ID NOS: 52 and 53), are Claudin molecules which have one or more of the following activities: (1) it elicits fluid accumulation in the intestinal tract by altering the membrane permeability of intestinal epithelial cells and (2) thus acts as the causative agent of diarrhea. The polypeptides, nucleic acids, and modulators thereof can be used to treat colonic disorders, such as congenital anomalies (*e.g.*, megacolon and imperforate anus), idiopathic disorders (*e.g.*, diverticular disease and melanosis coli), vascular lesions (*e.g.*, ischemic colistis, hemorrhoids, angiodysplasia), inflammatory diseases (*e.g.*, colitis (*e.g.*, idiopathic ulcerative colitis, pseudomembranous colitis), and lymphopathia venereum), Crohn's disease, and tumors (*e.g.*, hyperplastic polyps, adenomatous polyps, bronchogenic cancer, colonic carcinoma, squamous cell carcinoma, adenoacanthomas, sarcomas, lymphomas, argentaffinomas, carcinoids, and melanocarcinomas).

OV48 (SEQ ID NOS: 226 and 227), OV49 (SEQ ID NOS: 228 and 229) and OV50 (SEQ ID NOS: 230 and 231), markers for an osteopontin protein, have one or more of the following activities: (1) they act as a vessel extracellular matrix protein involved in calcification and (2) atherosclerosis. Hence, OV48, OV49 and OV50 nucleic acids, proteins, and modulators thereof can be used to treat heart disorders, *e.g.*, ischemic heart disease, atherosclerosis, hypertension, angina pectoris, Hypertrophic Cardiomyopathy, and congenital heart disease. They can also be used to treat

cardiovascular disorders, such as ischemic heart disease (*e.g.*, angina pectoris, myocardial infarction, and chronic ischemic heart disease), hypertensive heart disease, pulmonary heart disease, valvular heart disease (*e.g.*, rheumatic fever and rheumatic heart disease, endocarditis, mitral valve prolapse, and aortic valve stenosis), congenital heart disease (*e.g.*, valvular and vascular obstructive lesions, atrial or ventricular septal defect, and patent ductus arteriosus), or myocardial disease (*e.g.*, myocarditis, congestive cardiomyopathy, and hypertrophic cariomyopathy).

OV37 (SEQ ID NOS: 176 and 177), a lipocalin marker, is known to be a component of the neutrophil gelatinase complex. OV37 nucleic acids, proteins, and modulators thereof can be used to modulate the proliferation, differentiation, and/or function of leukocytes. Thus, OV37 nucleic acids, proteins, and modulators thereof can be used to treat bone marrow, blood, and hematopoietic associated diseases and disorders, *e.g.*, acute myeloid leukemia, hemophilia, leukemia, anemia (*e.g.*, sickle cell anemia), and thalassemia. OV37 polypeptides, nucleic acids, and modulators thereof can be used to treat leukocytic disorders, such as leukopenias (*e.g.*, neutropenia, monocytopenia, lymphopenia, and granulocytopenia), leukocytosis (*e.g.*, granulocytosis, lymphocytosis, eosinophilia, monocytosis, acute and chronic lymphadenitis), malignant lymphomas (*e.g.*, Non-Hodgkin's lymphomas, Hodgkin's lymphomas, leukemias, agnogenic myeloid metaplasia, multiple myeloma, plasmacytoma, Waldenstrom's macroglobulinemia, heavy-chain disease, monoclonal gammopathy, histiocytoses, eosinophilic granuloma, and angioimmunoblastic lymphadenopathy).

OV2 (SEQ ID NOS: 285 and 286), is known to be a protease inhibitor, which is associated with emphysema and liver disease. Hence OV2 polypeptides, nucleic acids, and modulators thereof can be used to diagnose and treat pulmonary (lung) disorders, such as atelectasis, cystic fibrosis, rheumatoid lung disease, pulmonary congestion or edema, chronic obstructive airway disease (*e.g.*, emphysema, chronic bronchitis, bronchial asthma, and bronchiectasis), diffuse interstitial diseases (*e.g.*, sarcoidosis, pneumoconiosis, hypersensitivity pneumonitis, bronchiolitis, Goodpasture's syndrome, idiopathic pulmonary fibrosis, idiopathic pulmonary hemosiderosis, pulmonary alveolar proteinosis, desquamative interstitial pneumonitis, chronic interstitial pneumonia, fibrosing alveolitis, hamman-rich syndrome, pulmonary eosinophilia, diffuse interstitial fibrosis, Wegener's granulomatosis, lymphomatoid

granulomatosis, and lipid pneumonia), or tumors (*e.g.*, bronchogenic carcinoma, bronchiolovlveolar carcinoma, bronchial carcinoid, hamartoma, and mesenchymal tumors). In another example, OV2 polypeptides, nucleic acids, and modulators thereof can be used to diagnose and treat hepatic (liver) disorders, such as jaundice, hepatic failure, hereditary hyperbiliruinemias (*e.g.*, Gilbert's syndrome, Crigler-Najjar syndromes and Dubin-Johnson and Rotor's syndromes), hepatic circulatory disorders (*e.g.*, hepatic vein thrombosis and portal vein obstruction and thrombosis), hepatitis (*e.g.*, chronic active hepatitis, acute viral hepatitis, and toxic and drug-induced hepatitis), cirrhosis (*e.g.*, alcoholic cirrhosis, biliary cirrhosis, and hemochromatosis), or malignant tumors (*e.g.*, primary carcinoma, hepatoma, hepatoblastoma, liver cysts, and angiosarcoma).

OV32 (SEQ ID NOS: 166 and 167) and OV33 (SEQ ID NOS: 168 and 169), kallikrein markers, are useful in detection of primary mammary carcinomas, as well as primary ovarian cancers. Hence, OV32 and OV33 polypeptides, nucleic acids, and modulators thereof can be used to diagnose and treat ovarian disorders, such as ovarian endometriosis, non-neoplastic cysts (*e.g.*, follicular and luteal cysts and polycystic ovaries) and tumors (*e.g.*, carcinomas, tumors of surface epithelium, germ cell tumors, ovarian fibroma, sex cord-stromal tumors, and ovarian cancers (*e.g.*, metastatic carcinomas, and ovarian teratoma)).

OV68 (SEQ ID NOS: 192 and 193), OV69 (SEQ ID NOS: 194 and 195), OV70 (SEQ ID NOS: 196 and 197), OV71 (SEQ ID NOS: 198 and 199), OV72 (SEQ ID NOS: 200 and 201), OV41 (SEQ ID NOS: 202 and 203), OV42 (SEQ ID NOS: 204 and 205), OV43 (SEQ ID NOS: 206 and 205), OV44 (SEQ ID NOS: 207 and 208) and OV83 (SEQ ID NOS: 209 and 210), are all mesothelin markers, and have been found to play a role in cellular adhesion. The nucleic acids, proteins, and modulators thereof can be used to diagnose, treat and modulate disorders associated with adhesion and migration of cells, *e.g.*, platelet aggregation disorders (*e.g.*, Glanzmann's thromboasthemia, which is a bleeding disorder characterized by failure of platelet aggregation in response to cell stimuli), inflammatory disorders (*e.g.*, leukocyte adhesion deficiency, which is a disorder associated with impaired migration of neutrophils to sites of extravascular inflammation), disorders associated with abnormal tissue migration during embryo development, and tumor metastasis.

OV17 (SEQ ID NOS: 110 and 111), OV18 (SEQ ID NOS: 112 and 111), OV19 (SEQ ID NOS: 113 and 111), OV20 (SEQ ID NOS: 114 and 111), OV21 (SEQ ID NOS: 115 and 111) and OV22 (SEQ ID NOS: 116 and 117) are folate receptors, which are known to be markers of ovarian cancer. The nucleic acids, proteins, and modulators thereof can be used to diagnose, treat and modulate ovarian disorders (*e.g.*, ovarian cyst, ovarian fibroma, ovarian endometriosis, ovarian teratoma). Although these markers have been previously associated with ovarian cancer, the expression of such markers has not yet been identified in combination with the expression of other markers including those of the present invention. Such combination of markers will provide improved methods of diagnosing, characterizing, managing and treating ovarian cancer.

OV66 (SEQ ID NOS: 54 and 55), OV7 (SEQ ID NOS: 56 and 57), OV8 (SEQ ID NOS: 58 and 59) and OV81 (SEQ ID NOS: 60 and 61) are ceruloplasmin markers, known to encode a plasma metalloprotein that binds copper in the plasma. The nucleic acids, proteins, and modulators thereof can be used to diagnose, treat and modulate disorders in blood haemostasis and diseases caused by such an imbalance *e.g.*, (1) cardiovascular diseases or disorders, such as ischemic heart disease (*e.g.*, angina pectoris, myocardial infarction, and chronic ischemic heart disease), hypertensive heart disease, pulmonary heart disease, valvular heart disease (*e.g.*, rheumatic fever and rheumatic heart disease, endocarditis, mitral valve prolapse, and aortic valve stenosis), congenital heart disease (*e.g.*, valvular and vascular obstructive lesions, atrial or ventricular septal defect, and patent ductus arteriosus), or myocardial disease (*e.g.*, myocarditis, congestive cardiomyopathy, and hypertrophic cariomyopathy); (2) neuronal diseases such as Alzheimers Disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers (*e.g.*, metastatic carcinoma of the brain, glioblastoma, lymphoma, astrocytoma, acoustic neuroma), hydrocephalus, and encephalitis; and (3) Wilson's Disease.

TABLE 1

Marker	Gene Name	SEQ ID NO (nts)	SEQ ID NO (AAs)	CDS
OV1	ABCB1: ATP-binding cassette, sub-family B (MDR/TAP), member 1	1	2	425..4264
M430	ADPRT: ADP-ribosyltransferase	3	4	160..3204
M571	ANXA2: annexin A2, variant 1	5	6	134..1153
M572	ANXA2: annexin A2, variant 2	7	8	50..1069
M573	ANXA4: annexin A4	9	10	74..1039
OV3	AQP5: aquaporin 5	11	12	519..1316
M352	ARHGAP8: Rho GTPase activating protein 8, variant 1	13	14	142..1536
M353	ARHGAP8: Rho GTPase activating protein 8, variant 2	15	16	1..2043
M354	ARHGAP8: Rho GTPase activating protein 8, variant 3	17	18	1..2256
M608	ARHGAP8: Rho GTPase activating protein 8, variant 4	17	19	1..2157
M355	ARHGAP8: Rho GTPase activating protein 8, variant 5	20	21	<1..1314
M356	ARHGAP8: Rho GTPase activating protein 8, variant 6	22	23	1..1902
M357	ARHGAP8: Rho GTPase activating protein 8, variant 7	24	25	<1..1281
M358	ARHGAP8: Rho GTPase activating protein 8, variant 8	26	27	1..1386
M359	ARHGAP8: Rho GTPase activating protein 8, variant 9	28	29	<1..1059
OV5	BICD1: Bicaudal D homolog 1 (Drosophila)	30	31	82..3009
M431	BTG2: BTG family, member 2	32	33	72..548
M432	CADPS: Ca2+-dependent activator protein for secretion	34	35	240..4412
M609	CDH1: cadherin 1, type 1, E-cadherin (epithelial)	36	37	125..2773
M433	CDH6: cadherin 6, type 2, K-cadherin	38	39	327..2699
M434	CDKN2A: cyclin-dependent kinase inhibitor 2A	40	41	41..511
OV9	CGN: cingulin	42	43	152..3763
OV6	CHI3L1: cartilage glycoprotein-39	44	45	127..1278
M435	CKMT1: creatine kinase, mitochondrial 1 (ubiquitous)	46	47	164..1417
M604	CLDN10: claudin 10	48	49	36..772
OV10	CLDN16: claudin 16	50	51	69..986
M360	CLDN4: claudin 4	52	53	183..812
OV66	CP: ceruloplasmin (ferroxidase), variant 1	54	55	1..3210
OV7	CP: ceruloplasmin (ferroxidase), variant 2	56	57	<1..2561
OV8	CP: ceruloplasmin (ferroxidase), variant 3	58	59	1..3198
OV81	CP: ceruloplasmin (ferroxidase), variant 4	60	61	76..3348
M103	CRABP2: cellular retinoic acid-binding protein 2	62	63	138..554

OV40	DD96: Epithelial protein up-regulated in carcinoma, membrane associated protein 17	64	65	202..546
OV4	DEC2: basic helix-loop-helix protein	66	67	135..1583
M575	dehydrogenase	68	69	339..1364
M436	DLX5: distal-less homeo box 5	70	71	204..1073
OV12	EAB1: Eab1 protein	72	73	<1..1305
OV13	ESX protein	74	75	96..1211
OV67	EVI-1: Evi-1 protein, variant 1	76	77	250..2406
OV14	EVI-1: Evi-1 protein, variant 2	78	79	250..3405
OV15	EVI-1: Evi-1 protein, variant 3	80	81	250..2433
OV16	EVI-1: Evi-1 protein, variant 4	82	83	250..3378
M437	FLJ10546: hypothetical protein FLJ10546	84	85	28..1815
OV28	FLJ12799: hypothetical protein FLJ12799	86	87	39..797
M576	FLJ13710: hypothetical protein FLJ13710	88	89	96..1712
M438	FLJ13782: hypothetical protein FLJ13782	90	91	13..1890
OV29	FLJ20150: hypothetical protein FLJ20150	92	93	78..983
M439	FLJ20327: hypothetical protein FLJ20327	94	95	306..2186
M440	FLJ20758: hypothetical protein FLJ20758, variant 1	96	97	<2..1270
M441	FLJ20758: hypothetical protein FLJ20758, variant 2	98	99	<2..2095
M442	FLJ20758: hypothetical protein FLJ20758, variant 3	100	101	465..1307
M443	FLJ22252: likely ortholog of mouse SRY-box containing gene 17	102	103	205..1449
M444	FLJ22316: hypothetical protein FLJ22316	104	105	508..1206
M400	FLJ22418: hypothetical protein FLJ22418	106	107	71..919
M445	FLJ23499: hypothetical protein FLJ23499	108	109	21..473
OV17	FOLR1: folate receptor 1 (alpha), variant 1	110	111	139..912
OV18	FOLR1: folate receptor 1 (alpha), variant 2	112	111	211..984
OV19	FOLR1: folate receptor 1 (alpha), variant 3	113	111	46..819
OV20	FOLR1: folate receptor 1 (alpha), variant 4	114	111	437..1210
OV21	FOLR1: folate receptor 1 (alpha), variant 5	115	111	11..784
OV22	FOLR3: folate receptor 3 (gamma)	116	117	57..788
OV23	GPR39: G protein-coupled receptor 39	118	119	1..1362
M446	GPRC5B: G protein-coupled receptor, family C, group 5, member B	120	121	109..1320
OV24	G-protein coupled receptor	122	123	274..1236
M447	GRB7: growth factor receptor-bound protein 7	124	125	220..1818
OV11	HAIK1: type I intermediate filament cytokeratin	126	127	61..1329
M448	HOXB7: homeo box B7	128	129	100..753
M138	HSECP1: secretory protein, variant 1	130	131	27..863
M449	HSECP1: secretory protein, variant 2	132	133	136..768
M450	HSECP1: secretory protein, variant 3	134	135	202..933
M451	HSNFRK: HSNFRK protein	136	137	642..2939
OV26	hypothetical protein (1)	138	139	<1..1140
OV27	hypothetical protein (2)	140	141	242..1483
OV31	IFI30: interferon, gamma-inducible protein 30	142	143	41..952
OV58	IGF2: somatomedin A	144	145	553..1095

M452	IMP-2: IGF-II mRNA-binding protein 2	146	147	436..2106
M453	INDO: indoleamine-pyrrole 2, 3 dioxygenase	148	149	23..1234
OV73	IPT: tRNA isopentenylpyrophosphate transferase, variant 1	150	151	15..1418
M810	IPT: tRNA isopentenylpyrophosphate transferase, variant 2	152	153	15..1418
M454	ITGA3: integrin, alpha 3	154	155	74..3274
OV30	ITGB8: integrin, beta 8	156	157	681..2990
OV34	KIAA0762: KIAA0762 protein	158	159	<1..1875
M455	KIAA0869: KIAA0869 protein	160	161	<1..2668
OV35	KIAA1154: KIAA1154 protein	162	163	<1..677
OV36	KIAA1456: KIAA1456 protein	164	165	<366..1631
OV32	KLK10: kallikrein 10	166	167	82..912
OV33	KLK8: kallikrein 6	168	169	246..980
M456	KRT7: keratin 7, variant 1	170	171	57..1466
M611	KRT7: keratin 7, variant 2	172	173	54..1463
OV53	LC27: Putative integral membrane transporter	174	175	204..1055
OV37	LCN2: Lipocalin 2 (oncogene 24p3)	176	177	1..597
M457	LEFTB: left-right determination, factor B	178	179	71..1171
M559	LPHB: lipophilin B (uteroglobin family member), prostatein-like	180	181	64..336
OV38	LYST-interacting protein LIP6	182	183	11..586
OV39	MEIS1: MEIS1 protein	184	185	66..1238
M458	MGB2: mammaglobin 2	186	187	65..352
M459	MGC3184: similar to sialyltransferase 7 ((alpha-N-acetylneuraminyl 2, 3-betagalactosyl-1, 3)-N-acetyl galactosaminide alpha-2, 6-sialyltransferase) E	188	189	176..1186
OV52	MMP7: Matrix metalloproteinase 7 (matrilysin, uterine)	190	191	28..831
OV68	MSLN: mesothelin, variant 1	192	193	88..2196
OV69	MSLN: mesothelin, variant 2	194	195	88..1980
OV70	MSLN: mesothelin, variant 3	196	197	88..1950
OV71	MSLN: mesothelin, variant 4	198	199	88..2172
OV72	MSLN: mesothelin, variant 5	200	201	88..1926
OV41	MSLN: mesothelin, variant 6	202	203	<1..>1195
OV42	MSLN: mesothelin, variant 7	204	205	85..1953
OV43	MSLN: mesothelin, variant 8	206	205	88..1956
OV44	MSLN: mesothelin, variant 9	207	208	89..1975
OV83	MSLN: mesothelin, variant 10	209	210	295..2187
OV45	MUC1: mucin 1	211	212	58..1605
M460	MUC16: mucin 16, variant 1	213	214	<1..5352
M461	MUC16: mucin 16, variant 2	215	216	25..3471
M612	MUC16: mucin 16, variant 3	215	217	<1..5673
M462	MYOM2: myomesin (M-protein)	218	219	49..4448
M463	NaPi-lib: sodium dependent phosphate transporter isoform	220	221	36..2105
M464	NME5: protein expressed in non-metastatic cells 5	222	223	15..653

OV47	NUFIP1: nuclear fragile X mental retardation protein interacting protein 1	224	225	1..1488
OV48	OPN-a: Secreted phosphoprotein-1 (osteopontin, bone sialoprotein)	226	227	1..942
OV49	OPN-b: Secreted phosphoprotein-1 (osteopontin, bone sialoprotein)	228	229	88..990
OV50	OPN-c: Secreted phosphoprotein-1 (osteopontin, bone sialoprotein)	230	231	1..861
M578	PAEP: progestagen-associated endometrial protein, variant 1	232	233	36..578
M579	PAEP: progestagen-associated endometrial protein, variant 2	234	233	36..578
M580	PAEP: progestagen-associated endometrial protein, variant 3	235	233	36..578
M581	PAEP: progestagen-associated endometrial protein, variant 4	236	233	36..578
M583	PAEP: progestagen-associated endometrial protein, variant 5	237	238	45..305
M582	PAEP: progestagen-associated endometrial protein, variant 6	239	240	45..521
M613	PAEP: progestagen-associated endometrial protein, variant 7	239	241	45..521
M465	PAX8: paired box gene 8, isoform 8A	242	243	11..1363
M466	PAX8: paired box gene 8, isoform 8B, variant 1	244	245	11..1174
M614	PAX8: paired box gene 8, isoform 8B, variant 2	244	246	11..1174
M467	PAX8: paired box gene 8, isoform 8C	247	248	161..1357
M468	PAX8: paired box gene 8, isoform 8D	249	250	161..1126
M469	PAX8: paired box gene 8, isoform 8E	251	252	161..1024
M470	PRAME: preferentially expressed antigen in melanoma	253	254	236..1765
M615	PRKCI: protein kinase C, iota	255	256	205..1968
M605	PRP4: serine/threonine-protein kinase PRP4 homolog, variant 1	257	258	<1..3133
M606	PRP4: serine/threonine-protein kinase PRP4 homolog, variant 2	259	258	<1..3133
M607	PRP4: serine/threonine-protein kinase PRP4 homolog, variant 3	260	258	<1..3133
OV80	PRSS8: prostasin	261	262	229..1260
OV51	PTGS1: prostaglandin-endoperoxide synthase 1	263	264	6..1805
M312	PTK9: protein tyrosine kinase 9	265	266	61..1113
OV54	pyruvate dehydrogenase complex component E2	267	268	49..>358
OV55	S100A1: S100 calcium-binding protein A1	269	270	114..398
M471	S100A11: S100 calcium-binding protein A11 (calgizzarin)	271	272	121..438
M68	S100A2: S100 calcium-binding protein A2	273	274	41..334
M585	S100A6: S100 calcium-binding protein A6 (calcyclin)	275	276	103..375

OV57	SCNN1A: sodium channel, nonvoltage-gated 1 alpha, variant 1	277	278	100..2109
OV85	SCNN1A: sodium channel, nonvoltage-gated 1 alpha, variant 2	279	280	96..2105
M472	secreted protein (HETKL27)	281	282	88..618
M473	SEMA3A: sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3A	283	284	16..2331
OV2	SERPINA1: alpha-1 antitrypsin	285	286	35..1291
M474	Similar to hypothetical protein, MGC: 7199	287	288	173..1053
M586	Similar to proteasome (prosome, macropain) subunit, alpha type, 3	289	290	45..791
M587	Similar to zinc finger protein 136	291	292	139..1524
M475	SLPI: secretory leukocyte protease inhibitor (antileukoproteinase), variant 1	293	294	271..447
M185	SLPI: secretory leukocyte protease inhibitor (antileukoproteinase), variant 2	295	296	19..417
OV60	SNCG: synuclein, gamma	297	298	49..432
OV59	SORL1: sortilin-related receptor	299	300	198..6842
OV56	SPINT2: serine protease inhibitor, Kunitz type, 2, variant 1	301	302	301..1059
OV84	SPINT2: serine protease inhibitor, Kunitz type, 2, variant 2	303	304	332..919
OV65	SPON1: VSGP/F-spondin, variant 1	305	306	25..2448
M593	SPON1: VSGP/F-spondin, variant 2	307	308	180..2984
M594	SPON1: VSGP/F-spondin, variant 3	309	310	180..2687
OV82	ST14: matriptase	311	312	209..2557
M476	TACSTD2: tumor-associated calcium signal transducer 2	313	314	616..1587
M588	TFPI2: tissue factor pathway inhibitor 2	315	316	57..764
OV86	TMPRSS4: transmembrane protease, serine 4	317	318	310..1623
OV74	TPH: tryptophan hydroxylase, variant 1	319	320	1..1335
OV75	TPH: tryptophan hydroxylase, variant 2	321	322	1..1401
M327	TSPAN-1: Tetraspan NET-1 protein, variant 1	323	324	124..900
M328	TSPAN-1: Tetraspan NET-1 protein, variant 2	325	326	1..726
OV46	TTID: myotilin	327	328	281..1777
M589	UCH2: Ubiquitin carboxyl-terminal hydrolases family 2	329	330	551..2940
OV63	unnamed gene (1)	331	332	71..919
OV64	unnamed gene (2)	333	334	28..804
OV76	unnamed gene (3)	335	336	69..773
OV77	unnamed gene (4)	337	338	223..1284
OV78	unnamed gene (5), variant 1	339	340	84..2450
M616	unnamed gene (5), variant 2	341	342	84..2450
OV79	unnamed gene (6)	343	344	69..392
OV87	unnamed gene (7)	345	346	509..2428
OV88	unnamed gene (8)	347	348	71..919
M477	unnamed gene (9), variant 1	349	350	246..992
M617	unnamed gene (9), variant 2	349	351	246..992
M478	unnamed gene (9), variant 3	352	353	246..1004

M479	unnamed gene (9), variant 4	354	355	246..1049
M590	unnamed gene (10), variant 1	356	357	21..404
M591	unnamed gene (10), variant 2	358	357	21..404
M592	unnamed gene (10), variant 3	359	357	21..404
OV25	WFDC2: Epididymis-specific, whey-acidic protein type, four-disulfide core; putative ovarian carcinoma marker	360	361	28..405
M480	XRCC5, KU80: ATP-dependant DNA helicase II	362	363	34..2232

TABLE 2

Marker	Gene Name	SEQ ID NO (nts)	SEQ ID NO (AAs)	CDS
M354	ARHGAP8: Rho GTPase activating protein 8, variant 3	17	18	1..2256
M608	ARHGAP8: Rho GTPase activating protein 8, variant 4	17	19	1..2157
M355	ARHGAP8: Rho GTPase activating protein 8, variant 5	20	21	<1..1314
M356	ARHGAP8: Rho GTPase activating protein 8, variant 6	22	23	1..1902
M357	ARHGAP8: Rho GTPase activating protein 8, variant 7	24	25	<1..1281
M358	ARHGAP8: Rho GTPase activating protein 8, variant 8	26	27	1..1386
M359	ARHGAP8: Rho GTPase activating protein 8, variant 9	28	29	<1..1059
OV66	CP: ceruloplasmin (ferroxidase), variant 1	54	55	1..3210
OV81	CP: ceruloplasmin (ferroxidase), variant 4	60	61	76..3348
M575	dehydrogenase	68	69	339..1364
OV67	EVI-1: Evi-1 protein, variant 1	76	77	250..2406
M440	FLJ20758: hypothetical protein FLJ20758, variant 1	96	97	<2..1270
M441	FLJ20758: hypothetical protein FLJ20758, variant 2	98	99	<2..2095
M449	HSECP1: secretory protein, variant 2	132	133	136..768
M450	HSECP1: secretory protein, variant 3	134	135	202..933
OV73	IPT: tRNA isopentenylpyrophosphate transferase, variant 1	150	151	15..1418
M610	IPT: tRNA isopentenylpyrophosphate transferase, variant 2	152	153	15..1418
M611	KRT7: keratin 7, variant 2	172	173	54..1463
OV68	MSLN: mesothelin, variant 1	192	193	88..2196
OV69	MSLN: mesothelin, variant 2	194	195	88..1980
OV70	MSLN: mesothelin, variant 3	196	197	88..1950
OV71	MSLN: mesothelin, variant 4	198	199	88..2172
OV72	MSLN: mesothelin, variant 5	200	201	88..1926
OV83	MSLN: mesothelin, variant 10	209	210	295..2187
M460	MUC16: mucin 16, variant 1	213	214	<1..5352
M583	PAEP: progestagen-associated endometrial protein, variant 5	237	238	45..305

M613	PAEP: progestagen-associated endometrial protein, variant 7	239	241	45..521
M614	PAX8: paired box gene 8, isoform 8B, variant 2	244	246	11..1174
M605	PRP4: serine/threonine-protein kinase PRP4 homolog, variant 1	257	258	<1..3133
M606	PRP4: serine/threonine-protein kinase PRP4 homolog, variant 2	259	258	<1..3133
M607	PRP4: serine/threonine-protein kinase PRP4 homolog, variant 3	260	258	<1..3133
OV85	SCNN1A: sodium channel, nonvoltage-gated 1 alpha, variant 2	279	280	96..2105
M475	SLPI: secretory leukocyte protease inhibitor (antileukoproteinase), variant 1	293	294	271..447
OV84	SPINT2: serine protease inhibitor, Kunitz type, 2, variant 2	303	304	332..919
M593	SPON1: VSGP/F-spondin, variant 2	307	308	180..2984
M594	SPON1: VSGP/F-spondin, variant 3	309	310	180..2687
OV82	ST14: matriptase	311	312	209..2557
OV86	TMPRSS4: transmembrane protease, serine 4	317	318	310..1623
OV74	TPH: tryptophan hydroxylase, variant 1	319	320	1..1335
OV75	TPH: tryptophan hydroxylase, variant 2	321	322	1..1401
M327	TSPAN-1: Tetraspan NET-1 protein, variant 1	323	324	124..900
M589	UCH2: Ubiquitin carboxyl-terminal hydrolases family 2	329	330	551..2940
OV76	unnamed gene (3)	335	336	69..773
OV77	unnamed gene (4)	337	338	223..1284
OV78	unnamed gene (5), variant 1	339	340	84..2450
M616	unnamed gene (5), variant 2	341	342	84..2450
OV79	unnamed gene (6)	343	344	69..392
OV87	unnamed gene (7)	345	346	509..2428
OV88	unnamed gene (8)	347	348	71..919
M477	unnamed gene (9), variant 1	349	350	246..992
M617	unnamed gene (9), variant 2	349	351	246..992
M478	unnamed gene (9), variant 3	352	353	246..1004
M479	unnamed gene (9), variant 4	354	355	246..1049

TABLE 3

Marker	Gene Name	SEQ ID NO (nts)	SEQ ID NO (AAs)	CDS
M604	CLDN10: claudin 10	48	49	36..772
OV14	EVI-1: Evi-1 protein, variant 2	78	79	250..3405
OV15	EVI-1: Evi-1 protein, variant 3	80	81	250..2433
OV16	EVI-1: Evi-1 protein, variant 4	82	83	250..3378
M576	FLJ13710: hypothetical protein FLJ13710	88	89	96..1712
M444	FLJ22316: hypothetical protein FLJ22316	104	105	508..1206
OV30	ITGB8: integrin, beta 8	156	157	681..2990
OV43	MSLN: mesothelin, variant 8	206	205	88..1956

M581	PAEP: progestagen-associated endometrial protein, variant 4	236	233	36..578
M582	PAEP: progestagen-associated endometrial protein, variant 6	239	240	45..521
M466	PAX8: paired box gene 8, isoform 8B, variant 1	244	245	11..1174
M467	PAX8: paired box gene 8, isoform 8C	247	248	161..1357
M468	PAX8: paired box gene 8, isoform 8D	249	250	161..1126
M469	PAX8: paired box gene 8, isoform 8E	251	252	161..1024
OV2	SERPINA1: alpha-1 antitrypsin	285	286	35..1291
M474	Similar to hypothetical protein, MGC: 7199	287	288	173..1053
M590	unnamed gene (10), variant 1	356	357	21..404
M591	unnamed gene (10), variant 2	358	357	21..404
M592	unnamed gene (10), variant 3	359	357	21..404

Definitions

As used herein, each of the following terms has the meaning associated with it in this section.

The articles "a" and "an" are used herein to refer to one or to more than one (*i.e.* to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

A "marker" is a gene whose altered level of expression in a tissue or cell from its expression level in normal or healthy tissue or cell is associated with a disease state, such as cancer. A "marker nucleic acid" is a nucleic acid (*e.g.*, mRNA, cDNA) encoded by or corresponding to a marker of the invention. Such marker nucleic acids can be DNA (*e.g.*, cDNA) comprising the sequences listed in Table 1 or the complement of such sequences. The marker nucleic acids also can be RNA comprising the sequences listed in Table 1 or the complement of such sequence, wherein all thymidine residues are replaced with uridine residues. A "marker protein" is a protein encoded by or corresponding to a marker of the invention. A marker protein comprises the sequence of any of the sequences listed in Table 1. The terms "protein" and "polypeptide' are used interchangeably.

The term "probe" refers to any molecule which is capable of selectively binding to a specifically intended target molecule, for example, a nucleotide transcript or protein encoded by or corresponding to a marker. Probes can be either synthesized by one skilled in the art, or derived from appropriate biological preparations. For purposes of detection of the target molecule, probes may be specifically designed to be

labeled, as described herein. Examples of molecules that can be utilized as probes include, but are not limited to, RNA, DNA, proteins, antibodies, and organic molecules.

An "ovary-associated" body fluid is a fluid which, when in the body of a patient, contacts or passes through ovarian cells or into which cells or proteins shed from ovarian cells *e.g.*. ovarian epithelium, are capable of passing. Exemplary ovary-associated body fluids include blood fluids, lymph, ascites, gynecological fluids, cystic fluid, urine, and fluids collected by peritoneal rinsing.

The "normal" level of expression of a marker is the level of expression of the marker in ovarian cells of a human subject or patient not afflicted with ovarian cancer

An "over-expression" or "significantly higher level of expression" of a marker refers to an expression level in a test sample that is greater than the standard error of the assay employed to assess expression, and is preferably at least twice, and more preferably three, four, five or ten times the expression level of the marker in a control sample (*e.g.*, sample from a healthy subjects not having the marker associated disease) and preferably, the average expression level of the marker in several control samples.

As used herein, the term "promoter/regulatory sequence" means a nucleic acid sequence which is required for expression of a gene product operably linked to the promoter/regulatory sequence. In some instances, this sequence may be the core promoter sequence and in other instances, this sequence may also include an enhancer sequence and other regulatory elements which are required for expression of the gene product. The promoter/regulatory sequence may, for example, be one which expresses the gene product in a tissue-specific manner.

A "constitutive" promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a living human cell under most or all physiological conditions of the cell.

An "inducible" promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a living human cell substantially only when an inducer which corresponds to the promoter is present in the cell.

A "tissue-specific" promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a living human cell substantially only if the cell is a cell of the tissue type corresponding to the promoter.

A "transcribed polynucleotide" or "nucleotide transcript" is a polynucleotide (*e.g.* an mRNA, hnRNA, a cDNA, or an analog of such RNA or cDNA) which is complementary to or homologous with all or a portion of a mature mRNA made by transcription of a marker of the invention and normal post-transcriptional processing (*e.g.* splicing), if any, of the RNA transcript, and reverse transcription of the RNA transcript.

"Complementary" refers to the broad concept of sequence complementarity between regions of two nucleic acid strands or between two regions of the same nucleic acid strand. It is known that an adenine residue of a first nucleic acid region is capable of forming specific hydrogen bonds ("base pairing") with a residue of a second nucleic acid region which is antiparallel to the first region if the residue is thymine or uracil. Similarly, it is known that a cytosine residue of a first nucleic acid strand is capable of base pairing with a residue of a second nucleic acid strand which is antiparallel to the first strand if the residue is guanine. A first region of a nucleic acid is complementary to a second region of the same or a different nucleic acid if, when the two regions are arranged in an antiparallel fashion, at least one nucleotide residue of the first region is capable of base pairing with a residue of the second region. Preferably, the first region comprises a first portion and the second region comprises a second portion, whereby, when the first and second portions are arranged in an antiparallel fashion, at least about 50%, and preferably at least about 75%, at least about 90%, or at least about 95% of the nucleotide residues of the first portion are capable of base pairing with nucleotide residues in the second portion. More preferably, all nucleotide residues of the first portion are capable of base pairing with nucleotide residues in the second portion.

"Homologous" as used herein, refers to nucleotide sequence similarity between two regions of the same nucleic acid strand or between regions of two different nucleic acid strands. When a nucleotide residue position in both regions is occupied by the same nucleotide residue, then the regions are homologous at that position. A first

region is homologous to a second region if at least one nucleotide residue position of each region is occupied by the same residue. Homology between two regions is expressed in terms of the proportion of nucleotide residue positions of the two regions that are occupied by the same nucleotide residue. By way of example, a region having the nucleotide sequence 5'-ATTGCC-3' and a region having the nucleotide sequence 5'-TATGGC-3' share 50% homology. Preferably, the first region comprises a first portion and the second region comprises a second portion, whereby, at least about 50%, and preferably at least about 75%, at least about 90%, or at least about 95% of the nucleotide residue positions of each of the portions are occupied by the same nucleotide residue. More preferably, all nucleotide residue positions of each of the portions are occupied by the same nucleotide residue.

A molecule is "fixed" or "affixed" to a substrate if it is covalently or non-covalently associated with the substrate such the substrate can be rinsed with a fluid (*e.g.* standard saline citrate, pH 7.4) without a substantial fraction of the molecule dissociating from the substrate.

As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in an organism found in nature.

A cancer is "inhibited" if at least one symptom of the cancer is alleviated, terminated, slowed, or prevented. As used herein, ovarian cancer is also "inhibited" if recurrence or metastasis of the cancer is reduced, slowed, delayed, or prevented.

A kit is any manufacture (*e.g.* a package or container) comprising at least one reagent, *e.g.* a probe, for specifically detecting the expression of a marker of the invention. The kit may be promoted, distributed, or sold as a unit for performing the methods of the present invention.

"Proteins of the invention" encompass marker proteins and their fragments; variant marker proteins and their fragments; peptides and polypeptides comprising an at least 15 amino acid segment of a marker or variant marker protein; and fusion proteins comprising a marker or variant marker protein, or an at least 15 amino acid segment of a marker or variant marker protein.

Unless otherwise specified herewithin, the terms "antibody" and "antibodies" broadly encompass naturally-occurring forms of antibodies (*e.g.*, IgG, IgA, IgM, IgE) and recombinant antibodies such as single-chain antibodies, chimeric and humanized antibodies and multi-specific antibodies, as well as fragments and derivatives of all of the foregoing, which fragments and derivatives have at least an antigenic binding site. Antibody derivatives may comprise a protein or chemical moiety conjugated to an antibody moiety.

Description

The present invention is based, in part, on newly identified markers which are over-expressed in ovarian cancer cells as compared to their expression in normal (*i.e.* non-cancerous) ovarian cells. The enhanced expression of one or more of these markers in ovarian cells is herein correlated with the cancerous state of the tissue. The invention provides compositions, kits, and methods for assessing the cancerous state of ovarian cells (*e.g.* cells obtained from a human, cultured human cells, archived or preserved human cells and *in vivo* cells) as well as treating patients afflicted with ovarian cancer.

The compositions, kits, and methods of the invention have the following uses, among others:

1) assessing whether a patient is afflicted with ovarian cancer;
2) assessing the stage of ovarian cancer in a human patient;
3) assessing the grade of ovarian cancer in a patient;
4) assessing the benign or malignant nature of ovarian cancer in a patient;
5) assessing the metastatic potential of ovarian cancer in a patient;
6) assessing the histological type of neoplasm (*e.g.* serous, mucinous, endometroid, or clear cell neoplasm) associated with ovarian cancer in a patient;
7) making antibodies, antibody fragments or antibody derivatives that are useful for treating ovarian cancer and/or assessing whether a patient is afflicted with ovarian cancer;

8) assessing the presence of ovarian cancer cells;

9) assessing the efficacy of one or more test compounds for inhibiting ovarian cancer in a patient;

10) assessing the efficacy of a therapy for inhibiting ovarian cancer in a patient;

11) monitoring the progression of ovarian cancer in a patient;

12) selecting a composition or therapy for inhibiting ovarian cancer in a patient;

13) treating a patient afflicted with ovarian cancer;

14) inhibiting ovarian cancer in a patient;

15) assessing the ovarian carcinogenic potential of a test compound; and

16) preventing the onset of ovarian cancer in a patient at risk for developing ovarian cancer.

The invention thus includes a method of assessing whether a patient is afflicted with ovarian cancer which includes assessing whether the patient has pre-metastasized ovarian cancer. This method comprises comparing the level of expression of a marker of the invention (listed in Table 1) in a patient sample and the normal level of expression of the marker in a control, *e.g.*, a non-ovarian cancer sample. A significantly higher level of expression of the marker in the patient sample as compared to the normal level is an indication that the patient is afflicted with ovarian cancer.

Gene delivery vehicles, host cells and compositions (all described herein) containing nucleic acids comprising the entirety, or a segment of 15 or more nucleotides, of any of the sequences listed in Tables 1-3 or the complement of such sequences, and polypeptides comprising the entirety, or a segment of 10 or more amino acids, of any of the sequences listed in Tables 1-3 are also provided by this invention.

As described herein, ovarian cancer in patients is associated with an increased level of expression of one or more markers of the invention. While, as discussed above, some of these changes in expression level result from occurrence of the ovarian cancer, others of these changes induce, maintain, and promote the cancerous state of ovarian cancer cells. Thus, ovarian cancer characterized by an increase in the level of expression of one or more markers of the invention can be inhibited by reducing

and/or interfering with the expression of the markers and/or function of the proteins encoded by those markers.

Expression of a marker of the invention can be inhibited in a number of ways generally known in the art. For example, an antisense oligonucleotide can be provided to the ovarian cancer cells in order to inhibit transcription, translation, or both, of the marker(s). Alternately, a polynucleotide encoding an antibody, an antibody derivative, or an antibody fragment which specifically binds a marker protein, and operably linked with an appropriate promoter/regulator region, can be provided to the cell in order to generate intracellular antibodies which will inhibit the function or activity of the protein. The expression and/or function of a marker may also be inhibited by treating the ovarian cancer cell with an antibody, antibody derivative or antibody fragment that specifically binds a marker protein. Using the methods described herein, a variety of molecules, particularly including molecules sufficiently small that they are able to cross the cell membrane, can be screened in order to identify molecules which inhibit expression of a marker or inhibit the function of a marker protein. The compound so identified can be provided to the patient in order to inhibit ovarian cancer cells of the patient.

Any marker or combination of markers of the invention, as well as any known markers in combination with the markers of the invention, may be used in the compositions, kits, and methods of the present invention. In general, it is preferable to use markers for which the difference between the level of expression of the marker in ovarian cancer cells and the level of expression of the same marker in normal ovarian cells is as great as possible. Although this difference can be as small as the limit of detection of the method for assessing expression of the marker, it is preferred that the difference be at least greater than the standard error of the assessment method, and preferably a difference of at least 2-, 3-, 4-, 5-, 6-, 7-, 8-, 9-, 10-, 15-, 20-, 25-, 100-, 500-, 1000-fold or greater than the level of expression of the same marker in normal ovarian tissue.

It is recognized that certain marker proteins are secreted from ovarian cells (*i.e.* one or both of normal and cancerous cells) to the extracellular space surrounding the cells. These markers are preferably used in certain embodiments of the compositions, kits, and methods of the invention, owing to the fact that the such marker

proteins can be detected in an ovary-associated body fluid sample, which may be more easily collected from a human patient than a tissue biopsy sample. In addition, preferred *in vivo* techniques for detection of a marker protein include introducing into a subject a labeled antibody directed against the protein. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

It is a simple matter for the skilled artisan to determine whether any particular marker protein is a secreted protein. In order to make this determination, the marker protein is expressed in, for example, a mammalian cell, preferably a human ovarian cell line, extracellular fluid is collected, and the presence or absence of the protein in the extracellular fluid is assessed (*e.g.* using a labeled antibody which binds specifically with the protein).

The following is an example of a method which can be used to detect secretion of a protein. About 8 x 10^5 293T cells are incubated at 37°C in wells containing growth medium (Dulbecco's modified Eagle's medium {DMEM} supplemented with 10% fetal bovine serum) under a 5% (v/v) CO_2, 95% air atmosphere to about 60-70% confluence. The cells are then transfected using a standard transfection mixture comprising 2 micrograms of DNA comprising an expression vector encoding the protein and 10 microliters of LipofectAMINE™ (GIBCO/BRL Catalog no. 18342-012) per well. The transfection mixture is maintained for about 5 hours, and then replaced with fresh growth medium and maintained in an air atmosphere. Each well is gently rinsed twice with DMEM which does not contain methionine or cysteine (DMEM-MC; ICN Catalog no. 16-424- 54). About 1 milliliter of DMEM-MC and about 50 microcuries of Trans-^{35}S™ reagent (ICN Catalog no. 51006) are added to each well. The wells are maintained under the 5% CO_2 atmosphere described above and incubated at 37°C for a selected period. Following incubation, 150 microliters of conditioned medium is removed and centrifuged to remove floating cells and debris. The presence of the protein in the supernatant is an indication that the protein is secreted.

Examples of ovary-associated body fluids include blood fluids (*e.g.* whole blood, blood serum, blood having platelets removed therefrom, etc.), lymph, ascitic fluids, gynecological fluids (*e.g.* ovarian, fallopian, and uterine secretions, menses, vaginal douching fluids, fluids used to rinse ovarian cell samples, etc.), cystic fluid, urine, and fluids collected by peritoneal rinsing (*e.g.* fluids applied and collected during laparoscopy or fluids instilled into and withdrawn from the peritoneal cavity of a human patient). In these embodiments, the level of expression of the marker can be assessed by assessing the amount (*e.g.* absolute amount or concentration) of the marker protein in an ovary-associated body fluid obtained from a patient. The fluid can, of course, be subjected to a variety of well-known post-collection preparative and storage techniques (*e.g.* storage, freezing, ultrafiltration, concentration, evaporation, centrifugation, etc.) prior to assessing the amount of the marker in the fluid.

Many ovary-associated body fluids (*i.e.* usually excluding urine) can have ovarian cells, *e.g.* ovarian epithelium, therein, particularly when the ovarian cells are cancerous, and, more particularly, when the ovarian cancer is metastasizing. Cell-containing fluids which can contain ovarian cancer cells include, but are not limited to, peritoneal ascites, fluids collected by peritoneal rinsing, fluids collected by uterine rinsing, uterine fluids such as uterine exudate and menses, pleural fluid, and ovarian exudates. Thus, the compositions, kits, and methods of the invention can be used to detect expression of marker proteins having at least one portion which is displayed on the surface of cells which express it. It is a simple matter for the skilled artisan to determine whether a marker protein, or a portion thereof, is exposed on the cell surface. For example, immunological methods may be used to detect such proteins on whole cells, or well known computer-based sequence analysis methods (*e.g.* the SIGNALP program; Nielsen *et al.*, 1997, *Protein Engineering* 10:1-6) may be used to predict the presence of at least one extracellular domain (*i.e.* including both secreted proteins and proteins having at least one cell-surface domain). Expression of a marker protein having at least one portion which is displayed on the surface of a cell which expresses it may be detected without necessarily lysing the cell (*e.g.* using a labeled antibody which binds specifically with a cell-surface domain of the protein).

Expression of a marker of the invention may be assessed by any of a wide variety of well known methods for detecting expression of a transcribed nucleic acid or protein. Non-limiting examples of such methods include immunological methods for detection of secreted, cell-surface, cytoplasmic, or nuclear proteins, protein purification methods, protein function or activity assays, nucleic acid hybridization methods, nucleic acid reverse transcription methods, and nucleic acid amplification methods.

In a preferred embodiment, expression of a marker is assessed using an antibody (*e.g.* a radio-labeled, chromophore-labeled, fluorophore-labeled, or enzyme-labeled antibody), an antibody derivative (*e.g.* an antibody conjugated with a substrate or with the protein or ligand of a protein-ligand pair {*e.g.* biotin-streptavidin}), or an antibody fragment (*e.g.* a single-chain antibody, an isolated antibody hypervariable domain, etc.) or derivative which binds specifically with a marker protein or fragment thereof, including a marker protein which has undergone all or a portion of its normal post-translational modification.

In another preferred embodiment, expression of a marker is assessed by preparing mRNA/cDNA (*i.e.* a transcribed polynucleotide) from cells in a patient sample, and by hybridizing the mRNA/cDNA with a reference polynucleotide which is a complement of a marker nucleic acid, or a fragment thereof. cDNA can, optionally, be amplified using any of a variety of polymerase chain reaction methods prior to hybridization with the reference polynucleotide; preferably, it is not amplified. Expression of one or more markers can likewise be detected using quantitative PCR to assess the level of expression of the marker(s). Alternatively, any of the many known methods of detecting mutations or variants (*e.g.* single nucleotide polymorphisms, deletions, etc.) of a marker of the invention may be used to detect occurrence of a marker in a patient.

In a related embodiment, a mixture of transcribed polynucleotides obtained from the sample is contacted with a substrate having fixed thereto a polynucleotide complementary to or homologous with at least a portion (*e.g.* at least 7, 10, 15, 20, 25, 30, 40, 50, 100, 500, or more nucleotide residues) of a marker nucleic acid. If polynucleotides complementary to or homologous with several marker nucleic acids are differentially detectable on the substrate (*e.g.* detectable using different

chromophores or fluorophores, or fixed to different selected positions), then the levels of expression of a plurality of markers can be assessed simultaneously using a single substrate (*e.g.* a "gene chip" microarray of polynucleotides fixed at selected positions). When a method of assessing marker expression is used which involves hybridization of one nucleic acid with another, it is preferred that the hybridization be performed under stringent hybridization conditions.

Because the compositions, kits, and methods of the invention rely on detection of a difference in expression levels of one or more markers of the invention, it is preferable that the level of expression of the marker is significantly greater than the minimum detection limit of the method used to assess expression in at least one of normal ovarian cells and cancerous ovarian cells.

It is understood that by routine screening of additional patient samples using one or more of the markers of the invention, it will be realized that certain of the markers are over-expressed in cancers of various types, including specific ovarian cancers, as well as other cancers such as breast cancer, cervical cancer, etc. For example, it will be confirmed that some of the markers of the invention are over-expressed in most (*i.e.* 50% or more) or substantially all (*i.e.* 80% or more) of ovarian cancer. Furthermore, it will be confirmed that certain of the markers of the invention are associated with ovarian cancer of various stages (*i.e.* stage I, II, III, and IV ovarian cancers, as well as subclassifications IA, IB, IC, IIA, IIB, IIC, IIIA, IIIB, and IIIC, using the FIGO Stage Grouping system for primary carcinoma of the ovary; 1987, *Am. J. Obstet. Gynecol.* 156:236), of various histologic subtypes (*e.g.* serous, mucinous, endometroid, and clear cell subtypes, as well as subclassifications and alternate classifications adenocarcinoma, papillary adenocarcinoma, papillary cystadenocarcinoma, surface papillary carcinoma, malignant adenofibroma, cystadenofibroma, adenocarcinoma, cystadenocarcinoma, adenoacanthoma, endometrioid stromal sarcoma, mesodermal (Müllerian) mixed tumor, mesonephroid tumor, malignant carcinoma, Brenner tumor, mixed epithelial tumor, and undifferentiated carcinoma, using the WHO/FIGO system for classification of malignant ovarian tumors; Scully, *Atlas of Tumor Pathology*, 3d series, Washington DC), and various grades (*i.e.* grade I {well differentiated} , grade II {moderately well differentiated}, and grade III {poorly differentiated from surrounding normal tissue}).

In addition, as a greater number of patient samples are assessed for expression of the markers of the invention and the outcomes of the individual patients from whom the samples were obtained are correlated, it will also be confirmed that increased expression of certain of the markers of the invention are strongly correlated with malignant cancers and that increased expression of other markers of the invention are strongly correlated with benign tumors. The compositions, kits, and methods of the invention are thus useful for characterizing one or more of the stage, grade, histological type, and benign/malignant nature of ovarian cancer in patients. In addition, these compositions, kits, and methods can be used to detect and differentiate epithelial, stromal, and germ cell ovarian cancers.

When the compositions, kits, and methods of the invention are used for characterizing one or more of the stage, grade, histological type, and benign/malignant nature of ovarian cancer in a patient, it is preferred that the marker or panel of markers of the invention is selected such that a positive result is obtained in at least about 20%, and preferably at least about 40%, 60%, or 80%, and more preferably in substantially all patients afflicted with an ovarian cancer of the corresponding stage, grade, histological type, or benign/malignant nature. Preferably, the marker or panel of markers of the invention is selected such that a PPV of greater than about 10% is obtained for the general population (more preferably coupled with an assay specificity greater than 99.5%).

When a plurality of markers of the invention are used in the compositions, kits, and methods of the invention, the level of expression of each marker in a patient sample can be compared with the normal level of expression of each of the plurality of markers in non-cancerous samples of the same type, either in a single reaction mixture (*i.e.* using reagents, such as different fluorescent probes, for each marker) or in individual reaction mixtures corresponding to one or more of the markers. In one embodiment, a significantly increased level of expression of more than one of the plurality of markers in the sample, relative to the corresponding normal levels, is an indication that the patient is afflicted with ovarian cancer. When a plurality of markers is used, it is preferred that 2, 3, 4, 5, 8, 10, 12, 15, 20, 30, or 50 or more individual markers be used, wherein fewer markers are preferred.

In order to maximize the sensitivity of the compositions, kits, and methods of the invention (*i.e.* by interference attributable to cells of non-ovarian origin in a patient sample), it is preferable that the marker of the invention used therein be a marker which has a restricted tissue distribution, *e.g.*, normally not expressed in a non-epithelial tissue, and more preferably a marker which is normally not expressed in a non-ovarian tissue.

Only a small number of markers are known to be associated with ovarian cancers (*e.g. AKT2*, *Ki-RAS*, *ERBB2*, *c-MYC*, *RB1*, and *TP53*; Lynch, *supra*). These markers are not, of course, included among the markers of the invention, although they may be used together with one or more markers of the invention in a panel of markers, for example. It is well known that certain types of genes, such as oncogenes, tumor suppressor genes, growth factor-like genes, protease-like genes, and protein kinase-like genes are often involved with development of cancers of various types. Thus, among the markers of the invention, use of those which correspond to proteins which resemble proteins encoded by known oncogenes and tumor suppressor genes, and those which correspond to proteins which resemble growth factors, proteases, and protein kinases are preferred.

It is recognized that the compositions, kits, and methods of the invention will be of particular utility to patients having an enhanced risk of developing ovarian cancer and their medical advisors. Patients recognized as having an enhanced risk of developing ovarian cancer include, for example, patients having a familial history of ovarian cancer, patients identified as having a mutant oncogene (*i.e.* at least one allele), and patients of advancing age (*i.e.* women older than about 50 or 60 years).

The level of expression of a marker in normal (*i.e.* non-cancerous) human ovarian tissue can be assessed in a variety of ways. In one embodiment, this normal level of expression is assessed by assessing the level of expression of the marker in a portion of ovarian cells which appears to be non-cancerous and by comparing this normal level of expression with the level of expression in a portion of the ovarian cells which is suspected of being cancerous. For example, when laparoscopy or other medical procedure, reveals the presence of a lump on one portion of a patient's ovary, but not on another portion of the same ovary or on the other ovary, the normal level of expression of a marker may be assessed using one or both or the non-affected ovary and

a non-affected portion of the affected ovary, and this normal level of expression may be compared with the level of expression of the same marker in an affected portion (*i.e.* the lump) of the affected ovary. Alternately, and particularly as further information becomes available as a result of routine performance of the methods described herein, population-average values for normal expression of the markers of the invention may be used. In other embodiments, the 'normal' level of expression of a marker may be determined by assessing expression of the marker in a patient sample obtained from a non-cancer-afflicted patient, from a patient sample obtained from a patient before the suspected onset of ovarian cancer in the patient, from archived patient samples, and the like.

The invention includes compositions, kits, and methods for assessing the presence of ovarian cancer cells in a sample (*e.g.* an archived tissue sample or a sample obtained from a patient). These compositions, kits, and methods are substantially the same as those described above, except that, where necessary, the compositions, kits, and methods are adapted for use with samples other than patient samples. For example, when the sample to be used is a parafinized, archived human tissue sample, it can be necessary to adjust the ratio of compounds in the compositions of the invention, in the kits of the invention, or the methods used to assess levels of marker expression in the sample. Such methods are well known in the art and within the skill of the ordinary artisan.

The invention includes a kit for assessing the presence of ovarian cancer cells (*e.g.* in a sample such as a patient sample). The kit comprises a plurality of reagents, each of which is capable of binding specifically with a marker nucleic acid or protein. Suitable reagents for binding with a marker protein include antibodies, antibody derivatives, antibody fragments, and the like. Suitable reagents for binding with a marker nucleic acid (*e.g.* a genomic DNA, an mRNA, a spliced mRNA, a cDNA, or the like) include complementary nucleic acids. For example, the nucleic acid reagents may include oligonucleotides (labeled or non-labeled) fixed to a substrate, labeled oligonucleotides not bound with a substrate, pairs of PCR primers, molecular beacon probes, and the like.

The kit of the invention may optionally comprise additional components useful for performing the methods of the invention. By way of example, the kit may comprise fluids (*e.g.* SSC buffer) suitable for annealing complementary nucleic acids or for binding an antibody with a protein with which it specifically binds, one or more sample compartments, an instructional material which describes performance of a method of the invention, a sample of normal ovarian cells, a sample of ovarian cancer cells, and the like.

The invention also includes a method of making an isolated hybridoma which produces an antibody useful for assessing whether patient is afflicted with an ovarian cancer. In this method, a protein or peptide comprising the entirety or a segment of a marker protein is synthesized or isolated (*e.g.* by purification from a cell in which it is expressed or by transcription and translation of a nucleic acid encoding the protein or peptide *in vivo* or *in vitro* using known methods). A vertebrate, preferably a mammal such as a mouse, rat, rabbit, or sheep, is immunized using the protein or peptide. The vertebrate may optionally (and preferably) be immunized at least one additional time with the protein or peptide, so that the vertebrate exhibits a robust immune response to the protein or peptide. Splenocytes are isolated from the immunized vertebrate and fused with an immortalized cell line to form hybridomas, using any of a variety of methods well known in the art. Hybridomas formed in this manner are then screened using standard methods to identify one or more hybridomas which produce an antibody which specifically binds with the marker protein or a fragment thereof. The invention also includes hybridomas made by this method and antibodies made using such hybridomas.

The invention also includes a method of assessing the efficacy of a test compound for inhibiting ovarian cancer cells. As described above, differences in the level of expression of the markers of the invention correlate with the cancerous state of ovarian cells. Although it is recognized that changes in the levels of expression of certain of the markers of the invention likely result from the cancerous state of ovarian cells, it is likewise recognized that changes in the levels of expression of other of the markers of the invention induce, maintain, and promote the cancerous state of those cells. Thus, compounds which inhibit an ovarian cancer in a patient will cause the level of expression of one or more of the markers of the invention to change to a level nearer

the normal level of expression for that marker (*i.e.* the level of expression for the marker in non-cancerous ovarian cells).

This method thus comprises comparing expression of a marker in a first ovarian cell sample and maintained in the presence of the test compound and expression of the marker in a second ovarian cell sample and maintained in the absence of the test compound. A significantly reduced expression of a marker of the invention in the presence of the test compound is an indication that the test compound inhibits ovarian cancer. The ovarian cell samples may, for example, be aliquots of a single sample of normal ovarian cells obtained from a patient, pooled samples of normal ovarian cells obtained from a patient, cells of a normal ovarian cell line, aliquots of a single sample of ovarian cancer cells obtained from a patient, pooled samples of ovarian cancer cells obtained from a patient, cells of an ovarian cancer cell line, or the like. In one embodiment, the samples are ovarian cancer cells obtained from a patient and a plurality of compounds known to be effective for inhibiting various ovarian cancers are tested in order to identify the compound which is likely to best inhibit the ovarian cancer in the patient.

This method may likewise be used to assess the efficacy of a therapy for inhibiting ovarian cancer in a patient. In this method, the level of expression of one or more markers of the invention in a pair of samples (one subjected to the therapy, the other not subjected to the therapy) is assessed. As with the method of assessing the efficacy of test compounds, if the therapy induces a significantly lower level of expression of a marker of the invention then the therapy is efficacious for inhibiting ovarian cancer. As above, if samples from a selected patient are used in this method, then alternative therapies can be assessed *in vitro* in order to select a therapy most likely to be efficacious for inhibiting ovarian cancer in the patient.

As described above, the cancerous state of human ovarian cells is correlated with changes in the levels of expression of the markers of the invention. The invention includes a method for assessing the human ovarian cell carcinogenic potential of a test compound. This method comprises maintaining separate aliquots of human ovarian cells in the presence and absence of the test compound. Expression of a marker of the invention in each of the aliquots is compared. A significantly higher level of expression of a marker of the invention in the aliquot maintained in the presence of the

test compound (relative to the aliquot maintained in the absence of the test compound) is an indication that the test compound possesses human ovarian cell carcinogenic potential. The relative carcinogenic potentials of various test compounds can be assessed by comparing the degree of enhancement or inhibition of the level of expression of the relevant markers, by comparing the number of markers for which the level of expression is enhanced or inhibited, or by comparing both.

Various aspects of the invention are described in further detail in the following subsections.

I. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules, including nucleic acids which encode a marker protein or a portion thereof. Isolated nucleic acids of the invention also include nucleic acid molecules sufficient for use as hybridization probes to identify marker nucleic acid molecules, and fragments of marker nucleic acid molecules, *e.g.*, those suitable for use as PCR primers for the amplification or mutation of marker nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid molecule. Preferably, an "isolated" nucleic acid molecule is free of sequences (preferably protein-encoding sequences) which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kB, 4 kB, 3 kB, 2 kB, 1 kB, 0.5 kB or 0.1 kB of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques,

or substantially free of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention can be isolated using standard molecular biology techniques and the sequence information in the database records described herein. Using all or a portion of such nucleic acid sequences, nucleic acid molecules of the invention can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook *et al.*, ed., *Molecular Cloning: A Laboratory Manual, 2nd ed.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

A nucleic acid molecule of the invention can be amplified using cDNA, mRNA, or genomic DNA as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, nucleotides corresponding to all or a portion of a nucleic acid molecule of the invention can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which has a nucleotide sequence complementary to the nucleotide sequence of a marker nucleic acid or to the nucleotide sequence of a nucleic acid encoding a marker protein. A nucleic acid molecule which is complementary to a given nucleotide sequence is one which is sufficiently complementary to the given nucleotide sequence that it can hybridize to the given nucleotide sequence thereby forming a stable duplex.

Moreover, a nucleic acid molecule of the invention can comprise only a portion of a nucleic acid sequence, wherein the full length nucleic acid sequence comprises a marker nucleic acid or which encodes a marker protein. Such nucleic acids can be used, for example, as a probe or primer. The probe/primer typically is used as one or more substantially purified oligonucleotides. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 7, preferably about 15, more preferably about 25, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, or 400 or more consecutive nucleotides of a nucleic acid of the invention.

Probes based on the sequence of a nucleic acid molecule of the invention can be used to detect transcripts or genomic sequences corresponding to one or more markers of the invention. The probe comprises a label group attached thereto, *e.g.*, a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as part of a diagnostic test kit for identifying cells or tissues which mis-express the protein, such as by measuring levels of a nucleic acid molecule encoding the protein in a sample of cells from a subject, *e.g.*, detecting mRNA levels or determining whether a gene encoding the protein has been mutated or deleted.

The invention further encompasses nucleic acid molecules that differ, due to degeneracy of the genetic code, from the nucleotide sequence of nucleic acids encoding a marker protein and thus encode the same protein.

It will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequence can exist within a population (*e.g.*, the human population). Such genetic polymorphisms can exist among individuals within a population due to natural allelic variation. An allele is one of a group of genes which occur alternatively at a given genetic locus. In addition, it will be appreciated that DNA polymorphisms that affect RNA expression levels can also exist that may affect the overall expression level of that gene (*e.g.*, by affecting regulation or degradation).

As used herein, the phrase "allelic variant" refers to a nucleotide sequence which occurs at a given locus or to a polypeptide encoded by the nucleotide sequence.

As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding a polypeptide corresponding to a marker of the invention. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of a given gene. Alternative alleles can be identified by sequencing the gene of interest in a number of different individuals. This can be readily carried out by using hybridization probes to identify the same genetic locus in a variety of individuals. Any and all such nucleotide variations and resulting amino acid polymorphisms or variations that are the result of natural allelic variation and that do not alter the functional activity are intended to be within the scope of the invention.

In another embodiment, an isolated nucleic acid molecule of the invention is at least 7, 15, 20, 25, 30, 40, 60, 80, 100, 150, 200, 250, 300, 350, 400, 450, 550, 650, 700, 800, 900, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2400, 2600, 2800, 3000, 3500, 4000, 4500, or more nucleotides in length and hybridizes under stringent conditions to a marker nucleic acid or to a nucleic acid encoding a marker protein. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% (65%, 70%, preferably 75%) identical to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in sections 6.3.1-6.3.6 of *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989). A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 50-65°C.

In addition to naturally-occurring allelic variants of a nucleic acid molecule of the invention that can exist in the population, the skilled artisan will further appreciate that sequence changes can be introduced by mutation thereby leading to changes in the amino acid sequence of the encoded protein, without altering the biological activity of the protein encoded thereby. For example, one can make nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are not conserved or only semi-conserved among homologs of various species may be non-essential for activity and thus would be likely targets for alteration. Alternatively, amino acid residues that are conserved among the homologs of various species (*e.g.*, murine and human) may be essential for activity and thus would not be likely targets for alteration.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding a variant marker protein that contain changes in amino acid residues that are not essential for activity. Such variant marker proteins differ in amino acid sequence from the naturally-occurring marker proteins, yet retain biological activity. In one embodiment, such a variant marker protein has an amino acid sequence that is at

least about 40% identical, 50%, 60%, 70%, 80%, 90%, 95%, or 98% identical to the amino acid sequence of a marker protein.

An isolated nucleic acid molecule encoding a variant marker protein can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of marker nucleic acids, such that one or more amino acid residue substitutions, additions, or deletions are introduced into the encoded protein. Mutations can be introduced by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), non-polar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Alternatively, mutations can be introduced randomly along all or part of the coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for biological activity to identify mutants that retain activity. Following mutagenesis, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

The present invention encompasses antisense nucleic acid molecules, *i.e.*, molecules which are complementary to a sense nucleic acid of the invention, *e.g.*, complementary to the coding strand of a double-stranded marker cDNA molecule or complementary to a marker mRNA sequence. Accordingly, an antisense nucleic acid of the invention can hydrogen bond to (*i.e.* anneal with) a sense nucleic acid of the invention. The antisense nucleic acid can be complementary to an entire coding strand, or to only a portion thereof, *e.g.*, all or part of the protein coding region (or open reading frame). An antisense nucleic acid molecule can also be antisense to all or part of a non-coding region of the coding strand of a nucleotide sequence encoding a marker protein.

The non-coding regions ("5' and 3' untranslated regions") are the 5' and 3' sequences which flank the coding region and are not translated into amino acids.

An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been sub-cloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a marker protein to thereby inhibit expression of the marker, *e.g.*, by inhibiting transcription and/or translation. The

hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. Examples of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site or infusion of the antisense nucleic acid into an ovary-associated body fluid. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

An antisense nucleic acid molecule of the invention can be an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual α-units, the strands run parallel to each other (Gaultier *et al.*, 1987, *Nucleic Acids Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.*, 1987, *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.*, 1987, *FEBS Lett.* 215:327-330).

The invention also encompasses ribozymes. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes as described in Haselhoff and Gerlach, 1988, *Nature* 334:585-591) can be used to catalytically cleave mRNA transcripts to thereby inhibit translation of the protein encoded by the mRNA. A ribozyme having specificity for a nucleic acid molecule encoding a marker protein can be designed based upon the nucleotide sequence of a cDNA corresponding to the marker. For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved

(see Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742). Alternatively, an mRNA encoding a polypeptide of the invention can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules (see, *e.g.*, Bartel and Szostak, 1993, *Science* 261:1411-1418).

The invention also encompasses nucleic acid molecules which form triple helical structures. For example, expression of a marker of the invention can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the gene encoding the marker nucleic acid or protein (*e.g.*, the promoter and/or enhancer) to form triple helical structures that prevent transcription of the gene in target cells. See generally Helene (1991) *Anticancer Drug Des.* 6(6):569-84; Helene (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14(12):807-15.

In various embodiments, the nucleic acid molecules of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.*, 1996, *Bioorganic & Medicinal Chemistry* 4(1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996), *supra*; Perry-O'Keefe *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93:14670-675.

PNAs can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup (1996), *supra*; or as probes or primers for DNA sequence and hybridization (Hyrup, 1996, *supra*; Perry-O'Keefe *et al.*, 1996, *Proc. Natl. Acad. Sci. USA* 93:14670-675).

In another embodiment, PNAs can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated which can combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup, 1996, *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996), *supra,* and Finn *et al.* (1996) *Nucleic Acids Res.* 24(17):3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs. Compounds such as 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite can be used as a link between the PNA and the 5' end of DNA (Mag *et al.*, 1989, *Nucleic Acids Res.* 17:5973-88). PNA monomers are then coupled in a step-wise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.*, 1996, *Nucleic Acids Res.* 24(17):3357-63). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser *et al.*, 1975, *Bioorganic Med. Chem. Lett.* 5:1119-11124).

In other embodiments, the oligonucleotide can include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.*, 1987, *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. WO 88/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. WO 89/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (see, *e.g.*, Krol *et al.*, 1988, *Bio/Techniques* 6:958-976) or intercalating agents (see, *e.g.*, Zon, 1988, *Pharm. Res.* 5:539-549). To this end, the oligonucleotide can be conjugated to another molecule, *e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The invention also includes molecular beacon nucleic acids having at least one region which is complementary to a nucleic acid of the invention, such that the molecular beacon is useful for quantitating the presence of the nucleic acid of the invention in a sample. A "molecular beacon" nucleic acid is a nucleic acid comprising a pair of complementary regions and having a fluorophore and a fluorescent quencher associated therewith. The fluorophore and quencher are associated with different portions of the nucleic acid in such an orientation that when the complementary regions are annealed with one another, fluorescence of the fluorophore is quenched by the quencher. When the complementary regions of the nucleic acid are not annealed with one another, fluorescence of the fluorophore is quenched to a lesser degree. Molecular beacon nucleic acids are described, for example, in U.S. Patent 5,876,930.

II. Isolated Proteins and Antibodies

One aspect of the invention pertains to isolated marker proteins and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise antibodies directed against a marker protein or a fragment thereof. In one embodiment, the native marker protein can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, a protein or peptide comprising the whole or a segment of the marker protein is produced by recombinant DNA techniques. Alternative to recombinant expression, such protein or peptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free of chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. Thus, protein that is substantially free of cellular material includes preparations of protein having less than about 30%, 20%, 10%, or 5% (by dry weight) of heterologous protein (also referred to herein as a "contaminating protein"). When the protein or biologically active portion thereof is recombinantly produced, it is

also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, 10%, or 5% of the volume of the protein preparation. When the protein is produced by chemical synthesis, it is preferably substantially free of chemical precursors or other chemicals, *i.e.*, it is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. Accordingly such preparations of the protein have less than about 30%, 20%, 10%, 5% (by dry weight) of chemical precursors or compounds other than the polypeptide of interest.

Biologically active portions of a marker protein include polypeptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the marker protein, which include fewer amino acids than the full length protein, and exhibit at least one activity of the corresponding full-length protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the corresponding full-length protein. A biologically active portion of a marker protein of the invention can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length. Moreover, other biologically active portions, in which other regions of the marker protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of the native form of the marker protein.

Preferred marker proteins are encoded by nucleotide sequences comprising the sequences listed in Tables 1-3. Other useful proteins are substantially identical (*e.g.*, at least about 40%, preferably 50%, 60%, 70%, 80%, 90%, 95%, or 99%) to one of these sequences and retain the functional activity of the corresponding naturally-occurring marker protein yet differ in amino acid sequence due to natural allelic variation or mutagenesis.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences

is a function of the number of identical positions shared by the sequences (*i.e.*, % identity = # of identical positions/total # of positions (*e.g.*, overlapping positions) x100). In one embodiment the two sequences are the same length.

The determination of percent identity between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the BLASTN and BLASTX programs of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-410. BLAST nucleotide searches can be performed with the BLASTN program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to a nucleic acid molecules of the invention. BLAST protein searches can be performed with the BLASTP program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to a protein molecules of the invention. To obtain gapped alignments for comparison purposes, a newer version of the BLAST algorithm called Gapped BLAST can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402, which is able to perform gapped local alignments for the programs BLASTN, BLASTP and BLASTX. Alternatively, PSI-Blast can be used to perform an iterated search which detects distant relationships between molecules. When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (*e.g.*, BLASTX and BLASTN) can be used. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, (1988) *CABIOS* 4:11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used. Yet another useful algorithm for identifying regions of local sequence similarity and alignment is the FASTA algorithm as described in Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85:2444-2448. When using the FASTA algorithm for comparing nucleotide or amino acid sequences, a PAM120 weight residue table can, for example, be used with a *k*-tuple value of 2.

The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, only exact matches are counted.

The invention also provides chimeric or fusion proteins comprising a marker protein or a segment thereof. As used herein, a "chimeric protein" or "fusion protein" comprises all or part (preferably a biologically active part) of a marker protein operably linked to a heterologous polypeptide (*i.e.*, a polypeptide other than the marker protein). Within the fusion protein, the term "operably linked" is intended to indicate that the marker protein or segment thereof and the heterologous polypeptide are fused in-frame to each other. The heterologous polypeptide can be fused to the amino-terminus or the carboxyl-terminus of the marker protein or segment.

One useful fusion protein is a GST fusion protein in which a marker protein or segment is fused to the carboxyl terminus of GST sequences. Such fusion proteins can facilitate the purification of a recombinant polypeptide of the invention.

In another embodiment, the fusion protein contains a heterologous signal sequence at its amino terminus. For example, the native signal sequence of a marker protein can be removed and replaced with a signal sequence from another protein. For example, the gp67 secretory sequence of the baculovirus envelope protein can be used as a heterologous signal sequence (Ausubel *et al.*, ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, NY, 1992). Other examples of eukaryotic heterologous signal sequences include the secretory sequences of melittin and human placental alkaline phosphatase (Stratagene; La Jolla, California). In yet another example, useful prokaryotic heterologous signal sequences include the phoA secretory signal (Sambrook *et al., supra*) and the protein A secretory signal (Pharmacia Biotech; Piscataway, New Jersey).

In yet another embodiment, the fusion protein is an immunoglobulin fusion protein in which all or part of a marker protein is fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand (soluble or membrane-bound) and a protein on the surface of a cell (receptor), to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion protein can be used to affect the bioavailability of a

cognate ligand of a marker protein. Inhibition of ligand/receptor interaction can be useful therapeutically, both for treating proliferative and differentiative disorders and for modulating (*e.g.* promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies directed against a marker protein in a subject, to purify ligands and in screening assays to identify molecules which inhibit the interaction of the marker protein with ligands.

Chimeric and fusion proteins of the invention can be produced by standard recombinant DNA techniques. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and re-amplified to generate a chimeric gene sequence (see, *e.g.,* Ausubel *et al., supra*). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the polypeptide of the invention.

A signal sequence can be used to facilitate secretion and isolation of marker proteins. Signal sequences are typically characterized by a core of hydrophobic amino acids which are generally cleaved from the mature protein during secretion in one or more cleavage events. Such signal peptides contain processing sites that allow cleavage of the signal sequence from the mature proteins as they pass through the secretory pathway. Thus, the invention pertains to marker proteins, fusion proteins or segments thereof having a signal sequence, as well as to such proteins from which the signal sequence has been proteolytically cleaved (*i.e.*, the cleavage products). In one embodiment, a nucleic acid sequence encoding a signal sequence can be operably linked in an expression vector to a protein of interest, such as a marker protein or a segment thereof. The signal sequence directs secretion of the protein, such as from a eukaryotic host into which the expression vector is transformed, and the signal sequence is subsequently or concurrently cleaved. The protein can then be readily purified from the extracellular medium by art recognized methods. Alternatively, the signal sequence can

be linked to the protein of interest using a sequence which facilitates purification, such as with a GST domain.

The present invention also pertains to variants of the marker proteins. Such variants have an altered amino acid sequence which can function as either agonists (mimetics) or as antagonists. Variants can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation. An agonist can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of the protein. An antagonist of a protein can inhibit one or more of the activities of the naturally occurring form of the protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the protein of interest. Thus, specific biological effects can be elicited by treatment with a variant of limited function. Treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein can have fewer side effects in a subject relative to treatment with the naturally occurring form of the protein.

Variants of a marker protein which function as either agonists (mimetics) or as antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of the protein of the invention for agonist or antagonist activity. In one embodiment, a variegated library of variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential protein sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display). There are a variety of methods which can be used to produce libraries of potential variants of the marker proteins from a degenerate oligonucleotide sequence. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, 1983, *Tetrahedron* 39:3; Itakura *et al.*, 1984, *Annu. Rev. Biochem.* 53:323; Itakura *et al.*, 1984, *Science* 198:1056; Ike *et al.*, 1983 *Nucleic Acid Res.* 11:477).

In addition, libraries of segments of a marker protein can be used to generate a variegated population of polypeptides for screening and subsequent selection of variant marker proteins or segments thereof. For example, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of the

coding sequence of interest with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes amino terminal and internal fragments of various sizes of the protein of interest.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify variants of a protein of the invention (Arkin and Yourvan, 1992, *Proc. Natl. Acad. Sci. USA 89*:7811-7815; Delgrave *et al.*, 1993, *Protein Engineering* 6(3):327-331).

Another aspect of the invention pertains to antibodies directed against a protein of the invention. In preferred embodiments, the antibodies specifically bind a marker protein or a fragment thereof. The terms "antibody" and "antibodies" as used interchangeably herein refer to immunoglobulin molecules as well as fragments and derivatives thereof that comprise an immunologically active portion of an immunoglobulin molecule, (*i.e.*, such a portion contains an antigen binding site which specifically binds an antigen, such as a marker protein, *e.g.*, an epitope of a marker protein). An antibody which specifically binds to a protein of the invention is an antibody which binds the protein, but does not substantially bind other molecules in a sample, *e.g.*, a biological sample, which naturally contains the protein. Examples of an immunologically active portion of an immunoglobulin molecule include, but are not limited to, single-chain antibodies (scAb), F(ab) and $F(ab')_2$ fragments.

An isolated protein of the invention or a fragment thereof can be used as an immunogen to generate antibodies. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments for use as immunogens. The antigenic peptide of a protein of the invention comprises at least 8 (preferably 10, 15, 20, or 30 or more) amino acid residues of the amino acid sequence of one of the proteins of the invention, and encompasses at least one epitope of the protein such that an antibody raised against the peptide forms a specific immune complex with the protein. Preferred epitopes encompassed by the antigenic peptide are regions that are located on the surface of the protein, *e.g.*, hydrophilic regions. Hydrophobicity sequence analysis, hydrophilicity sequence analysis, or similar analyses can be used to identify hydrophilic regions. In preferred embodiments, an isolated marker protein or fragment thereof is used as an immunogen.

An immunogen typically is used to prepare antibodies by immunizing a suitable (*i.e.* immunocompetent) subject such as a rabbit, goat, mouse, or other mammal or vertebrate. An appropriate immunogenic preparation can contain, for example, recombinantly-expressed or chemically-synthesized protein or peptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or a similar immunostimulatory agent. Preferred immunogen compositions are those that contain no other human proteins such as, for example, immunogen compositions made using a non-human host cell for recombinant expression of a protein of the invention. In such a manner, the resulting antibody compositions have reduced or no binding of human proteins other than a protein of the invention.

The invention provides polyclonal and monoclonal antibodies. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope. Preferred polyclonal and monoclonal antibody compositions are ones that have been selected for antibodies directed against a protein of the invention. Particularly preferred polyclonal and monoclonal antibody preparations are ones that contain only antibodies directed against a marker protein or fragment thereof.

Polyclonal antibodies can be prepared by immunizing a suitable subject with a protein of the invention as an immunogen The antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized polypeptide.At an appropriate time after immunization, *e.g.*, when the specific antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies (mAb) by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497, the human B cell hybridoma technique (see Kozbor *et al.*, 1983, *Immunol. Today* 4:72), the EBV-hybridoma technique (see Cole *et al.*, pp. 77-96 In *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., 1985) or trioma techniques. The technology for producing hybridomas is well known (see generally *Current Protocols in Immunology*, Coligan *et al.* ed., John Wiley & Sons, New York, 1994). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind the polypeptide of interest, *e.g.*, using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal antibody directed against a protein of the invention can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with the polypeptide of interest. Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, U.S. Patent No. 5,223,409; PCT Publication No. WO 92/18619; PCT Publication No. WO 91/17271; PCT Publication No. WO 92/20791; PCT Publication No. WO 92/15679; PCT Publication No. WO 93/01288; PCT Publication No. WO 92/01047; PCT Publication No. WO 92/09690; PCT Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J.* 12:725-734.

The invention also provides recombinant antibodies that specifically bind a protein of the invention. In preferred embodiments, the recombinant antibodies specifically binds a marker protein or fragment thereof. Recombinant antibodies include, but are not limited to, chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, single-chain antibodies and multi-specific antibodies. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region. (See, *e.g.,* Cabilly et al., U.S. Patent No. 4,816,567; and Boss et al., U.S. Patent No. 4,816,397, which are incorporated herein by reference in their entirety.) Single-chain antibodies have an antigen binding site and consist of single polypeptides. They can be produced by techniques known in the art, for example using methods described in Ladner *et. al* U.S. Pat. No. 4,946,778 (which is incorporated herein by reference in its entirety); Bird *et al.*, (1988) *Science* 242:423-426; Whitlow *et al.*, (1991) *Methods in Enzymology* 2:1-9; Whitlow *et al.*, (1991) *Methods in Enzymology* 2:97-105; and Huston *et al.*, (1991) *Methods in Enzymology Molecular Design and Modeling: Concepts and Applications* 203:46-88. Multi-specific antibodies are antibody molecules having at least two antigen-binding sites that specifically bind different antigens. Such molecules can be produced by techniques known in the art, for example using methods described in Segal, U.S. Patent No. 4,676,980 (the disclosure of which is incorporated herein by reference in its entirety); Holliger et al., (1993) *Proc. Natl. Acad. Sci. USA* 90:6444-6448; *Whitlow et al.*, (1994) *Protein Eng.* 7:1017-1026 and U.S. Pat. No. 6,121,424.

Humanized antibodies are antibody molecules from non-human species having one or more complementarity determining regions (CDRs) from the non-human species and a framework region from a human immunoglobulin molecule. (See, *e.g.,* Queen, U.S. Patent No. 5,585,089, which is incorporated herein by reference in its entirety.) Humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT Publication No. WO 87/02671; European Patent Application 184,187; European Patent Application 171,496; European Patent Application 173,494; PCT Publication No. WO 86/01533; U.S. Patent No. 4,816,567; European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu

et al. (1987) *J. Immunol.* 139:3521- 3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Cancer Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559); Morrison (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *Bio/Techniques* 4:214; U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

More particularly, humanized antibodies can be produced, for example, using transgenic mice which are incapable of expressing endogenous immunoglobulin heavy and light chains genes, but which can express human heavy and light chain genes. The transgenic mice are immunized in the normal fashion with a selected antigen, *e.g.*, all or a portion of a polypeptide corresponding to a marker of the invention. Monoclonal antibodies directed against the antigen can be obtained using conventional hybridoma technology. The human immunoglobulin transgenes harbored by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus, using such a technique, it is possible to produce therapeutically useful IgG, IgA and IgE antibodies. For an overview of this technology for producing human antibodies, see Lonberg and Huszar (1995) *Int. Rev. Immunol.* 13:65-93). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, see, *e.g.*, U.S. Patent 5,625,126; U.S. Patent 5,633,425; U.S. Patent 5,569,825; U.S. Patent 5,661,016; and U.S. Patent 5,545,806. In addition, companies such as Abgenix, Inc. (Freemont, CA), can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

Completely human antibodies which recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody, *e.g.*, a murine antibody, is used to guide the selection of a completely human antibody recognizing the same epitope (Jespers *et al.*, 1994, *Bio/technology* 12:899-903).

The antibodies of the invention can be isolated after production (*e.g.*, from the blood or serum of the subject) or synthesis and further purified by well-known techniques. For example, IgG antibodies can be purified using protein A chromatography. Antibodies specific for a protein of the invention can be selected or

(*e.g.*, partially purified) or purified by, *e.g.*, affinity chromatography. For example, a recombinantly expressed and purified (or partially purified) protein of the invention is produced as described herein, and covalently or non-covalently coupled to a solid support such as, for example, a chromatography column. The column can then be used to affinity purify antibodies specific for the proteins of the invention from a sample containing antibodies directed against a large number of different epitopes, thereby generating a substantially purified antibody composition, *i.e.*, one that is substantially free of contaminating antibodies. By a substantially purified antibody composition is meant, in this context, that the antibody sample contains at most only 30% (by dry weight) of contaminating antibodies directed against epitopes other than those of the desired protein of the invention, and preferably at most 20%, yet more preferably at most 10%, and most preferably at most 5% (by dry weight) of the sample is contaminating antibodies. A purified antibody composition means that at least 99% of the antibodies in the composition are directed against the desired protein of the invention.

In a preferred embodiment, the substantially purified antibodies of the invention may specifically bind to a signal peptide, a secreted sequence, an extracellular domain, a transmembrane or a cytoplasmic domain or cytoplasmic membrane of a protein of the invention. In a particularly preferred embodiment, the substantially purified antibodies of the invention specifically bind to a secreted sequence or an extracellular domain of the amino acid sequences of a protein of the invention. In a more preferred embodiment, the substantially purified antibodies of the invention specifically bind to a secreted sequence or an extracellular domain of the amino acid sequences of a marker protein.

An antibody directed against a protein of the invention can be used to isolate the protein by standard techniques, such as affinity chromatography or immunoprecipitation. Moreover, such an antibody can be used to detect the marker protein or fragment thereof (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the level and pattern of expression of the marker. The antibodies can also be used diagnostically to monitor protein levels in tissues or body fluids (*e.g.* in an ovary-associated body fluid) as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by the

use of an antibody derivative, which comprises an antibody of the invention coupled to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

Antibodies of the invention may also be used as therapeutic agents in treating cancers. In a preferred embodiment, completely human antibodies of the invention are used for therapeutic treatment of human cancer patients, particularly those having an ovarian cancer. In another preferred embodiment, antibodies that bind specifically to a marker protein or fragment thereof are used for therapeutic treatment. Further, such therapeutic antibody may be an antibody derivative or immunotoxin comprising an antibody conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (*e.g.*, methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (*e.g.*, mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclothosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (*e.g.*, daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (*e.g.*, dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (*e.g.*, vincristine and vinblastine).

The conjugated antibodies of the invention can be used for modifying a given biological response, for the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as ribosome-inhibiting protein (see Better et al., U.S. Patent No. 6,146,631, the disclosure of which is incorporated herein in its entirety), abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, .alpha.-interferon, .beta.-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophase colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, *e.g.*, Arnon et al., "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in Monoclonal Antibodies And Cancer Therapy, Reisfeld et al. (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom et al., "Antibodies For Drug Delivery", in Controlled Drug Delivery (2nd Ed.), Robinson et al. (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera et al. (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin et al. (eds.), pp. 303-16 (Academic Press 1985), and Thorpe et al., "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", Immunol. Rev., 62:119-58 (1982).

Accordingly, in one aspect, the invention provides substantially purified antibodies, antibody fragments and derivatives, all of which specifically bind to a protein of the invention and preferably, a marker protein. In various embodiments, the substantially purified antibodies of the invention, or fragments or derivatives thereof, can be human, non-human, chimeric and/or humanized antibodies. In another aspect, the invention provides non-human antibodies, antibody fragments and derivatives, all of which specifically bind to a protein of the invention and preferably, a marker protein. Such non-human antibodies can be goat, mouse, sheep, horse, chicken, rabbit, or rat

antibodies. Alternatively, the non-human antibodies of the invention can be chimeric and/or humanized antibodies. In addition, the non-human antibodies of the invention can be polyclonal antibodies or monoclonal antibodies. In still a further aspect, the invention provides monoclonal antibodies, antibody fragments and derivatives, all of which specifically bind to a protein of the invention and preferably, a marker protein. The monoclonal antibodies can be human, humanized, chimeric and/or non-human antibodies.

The invention also provides a kit containing an antibody of the invention conjugated to a detectable substance, and instructions for use. Still another aspect of the invention is a pharmaceutical composition comprising an antibody of the invention and a pharmaceutically acceptable carrier. In preferred embodiments, the pharmaceutical composition contains an antibody of the invention, a therapeutic moiety, and a pharmaceutically acceptable carrier.

III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a marker protein (or a portion of such a protein). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors, namely expression vectors, are capable of directing the expression of genes to which they are operably linked. In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids (vectors). However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective

retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell. This means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operably linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel, *Methods in Enzymology: Gene Expression Technology* vol.185, Academic Press, San Diego, CA (1991). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cell and those which direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein.

The recombinant expression vectors of the invention can be designed for expression of a marker protein or a segment thereof in prokaryotic (*e.g.*, *E. coli*) or eukaryotic cells (*e.g.*, insect cells {using baculovirus expression vectors}, yeast cells or mammalian cells). Suitable host cells are discussed further in Goeddel, *supra*. Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a

protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson, 1988, *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, 1988, *Gene* 69:301-315) and pET 11d (Studier *et al.*, p. 60-89, In *Gene Expression Technology: Methods in Enzymology* vol.185, Academic Press, San Diego, CA, 1991). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a co-expressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, p. 119-128, In *Gene Expression Technology: Methods in Enzymology* vol. 185, Academic Press, San Diego, CA, 1990. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, 1992, *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari *et al.*, 1987, *EMBO J.* 6:229-234), pMFa (Kurjan and Herskowitz, 1982, *Cell* 30:933-943), pJRY88 (Schultz *et al.*, 1987, *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and pPicZ (Invitrogen Corp, San Diego, CA).

Alternatively, the expression vector is a baculovirus expression vector. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, Sf 9 cells) include the pAc series (Smith *et al.*, 1983, *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers, 1989, *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, 1987, *Nature* 329:840) and pMT2PC (Kaufman *et al.*, 1987, *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook *et al.*, *supra*.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.*, 1987, *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton, 1988, *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore, 1989, *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.*, 1983, *Cell* 33:729-740; Queen and Baltimore, 1983, *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle, 1989, *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund *et al.*, 1985, *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters

(Kessel and Gruss, 1990, *Science* 249:374-379) and the α-fetoprotein promoter (Camper and Tilghman, 1989, *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operably linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to the mRNA encoding a polypeptide of the invention. Regulatory sequences operably linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue-specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid, or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub *et al.*, 1986, *Trends in Genetics*, Vol. 1(1).

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic (*e.g.*, *E. coli*) or eukaryotic cell (*e.g.*, insect cells, yeast or mammalian cells).

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection,

lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (*supra*), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, for resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce a marker protein or a segment thereof. Accordingly, the invention further provides methods for producing a marker protein or a segment thereof using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of the invention (into which a recombinant expression vector encoding a marker protein or a segment thereof has been introduced) in a suitable medium such that the is produced. In another embodiment, the method further comprises isolating the a marker protein or a segment thereof from the medium or the host cell.

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which a sequences encoding a marker protein or a segment thereof have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous sequences encoding a marker protein of the invention have been introduced into their genome or homologous recombinant animals in which endogenous gene(s) encoding a marker protein have been altered. Such animals are useful for studying the function and/or activity of the marker protein and for identifying and/or evaluating modulators of marker protein. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human

primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, an "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing a nucleic acid encoding a marker protein into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the transgene to direct expression of the polypeptide of the invention to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, U.S. Patent No. 4,873,191 and in Hogan, *Manipulating the Mouse Embryo*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the transgene in its genome and/or expression of mRNA encoding the transgene in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying the transgene can further be bred to other transgenic animals carrying other transgenes.

To create an homologous recombinant animal, a vector is prepared which contains at least a portion of a gene encoding a marker protein into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the gene. In a preferred embodiment, the vector is designed such that, upon homologous recombination, the endogenous gene is functionally disrupted (*i.e.*, no longer encodes a

functional protein; also referred to as a "knock out" vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous protein). In the homologous recombination vector, the altered portion of the gene is flanked at its 5' and 3' ends by additional nucleic acid of the gene to allow for homologous recombination to occur between the exogenous gene carried by the vector and an endogenous gene in an embryonic stem cell. The additional flanking nucleic acid sequences are of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector (see, *e.g.*, Thomas and Capecchi, 1987, *Cell* 51:503 for a description of homologous recombination vectors). The vector is introduced into an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced gene has homologously recombined with the endogenous gene are selected (see, *e.g.*, Li *et al.*, 1992, *Cell* 69:915). The selected cells are then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras (see, *e.g.*, Bradley, *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, Ed., IRL, Oxford, 1987, pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Current Opinion in Bio/Technology* 2:823-829 and in PCT Publication NOS. WO 90/11354, WO 91/01140, WO 92/0968, and WO 93/04169.

In another embodiment, transgenic non-human animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, *e.g.*, Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.*, 1991, *Science* 251:1351-1355). If a *cre/loxP* recombinase system is used to regulate expression of the

transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut *et al.* (1997) *Nature* 385:810-813 and PCT Publication NOS. WO 97/07668 and WO 97/07669.

IV. Pharmaceutical Compositions

The nucleic acid molecules, polypeptides, and antibodies (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

The invention includes methods for preparing pharmaceutical compositions for modulating the expression or activity of a marker nucleic acid or protein . Such methods comprise formulating a pharmaceutically acceptable carrier with an agent which modulates expression or activity of a marker nucleic acid or protein. Such compositions can further include additional active agents. Thus, the invention further includes methods for preparing a pharmaceutical composition by formulating a pharmaceutically acceptable carrier with an agent which modulates expression or activity of a marker nucleic acid or protein and one or more additional active compounds.

The invention also provides methods (also referred to herein as "screening assays") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, peptoids, small molecules or other drugs) which (a) bind to the marker, or (b) have a modulatory (*e.g.*, stimulatory or inhibitory) effect on the activity of the marker or, more specifically, (c) have a modulatory effect on the interactions of the marker with one or more of its natural substrates (*e.g.*, peptide, protein, hormone, co-factor, or nucleic acid), or (d) have a modulatory effect on the expression of the marker. Such assays typically comprise a reaction between the marker and one or more assay components. The other components may be either the test compound itself, or a combination of test compound and a natural binding partner of the marker.

The test compounds of the present invention may be obtained from any available source, including systematic libraries of natural and/or synthetic compounds. Test compounds may also be obtained by any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; peptoid libraries (libraries of molecules having the functionalities of peptides, but with a novel, non-peptide backbone which are resistant to enzymatic degradation but which nevertheless remain bioactive; see, *e.g.*, Zuckermann *et al.*, 1994, *J. Med. Chem.* 37:2678-85); spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library and peptoid library approaches are limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, 1997, *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994). *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop *et al.* (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (*e.g.,* Houghten, 1992, *Biotechniques* 13:412-421), or on beads (Lam, 1991, *Nature* 354:82-84), chips (Fodor, 1993, *Nature* 364:555-556), bacteria and/or spores, (Ladner, USP 5,223,409), plasmids (Cull *et al*, 1992, *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith, 1990, *Science* 249:386-390; Devlin, 1990, *Science* 249:404-406; Cwirla *et al,* 1990, *Proc. Natl. Acad. Sci.* 87:6378-6382; Felici, 1991, *J. Mol. Biol.* 222:301-310; Ladner, *supra.*).

In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a protein encoded by or corresponding to a marker or biologically active portion thereof. In another embodiment, the invention provides assays for screening candidate or test compounds which bind to a protein encoded by or corresponding to a marker or biologically active portion thereof. Determining the ability of the test compound to directly bind to a protein can be accomplished, for example, by coupling the compound with a radioisotope or enzymatic label such that binding of the compound to the marker can be determined by detecting the labeled marker compound in a complex. For example, compounds (*e.g.,* marker substrates) can be labeled with ^{125}I, ^{35}S, ^{14}C, or ^{3}H, either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, assay components can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product.

In another embodiment, the invention provides assays for screening candidate or test compounds which modulate the expression of a marker or the activity of a protein encoded by or corresponding to a marker, or a biologically active portion thereof. In all likelihood, the protein encoded by or corresponding to the marker can, *in vivo*, interact with one or more molecules, such as but not limited to, peptides, proteins, hormones, cofactors and nucleic acids. For the purposes of this discussion, such cellular and extracellular molecules are referred to herein as "binding partners" or marker "substrate".

One necessary embodiment of the invention in order to facilitate such screening is the use of a protein encoded by or corresponding to marker to identify the protein's natural *in vivo* binding partners. There are many ways to accomplish this which are known to one skilled in the art. One example is the use of the marker protein as "bait protein" in a two-hybrid assay or three-hybrid assay (see, *e.g.,* U.S. Patent No. 5,283,317; Zervos *et al*, 1993, *Cell* 72:223-232; Madura *et al,* 1993, *J. Biol. Chem.* 268:12046-12054; Bartel *et al* ,1993, *Biotechniques* 14:920-924; Iwabuchi *et al*, 1993 *Oncogene* 8:1693-1696; Brent WO94/10300) in order to identify other proteins which bind to or interact with the marker (binding partners) and, therefore, are possibly involved in the natural function of the marker. Such marker binding partners are also likely to be involved in the propagation of signals by the marker protein or downstream elements of a marker protein-mediated signaling pathway. Alternatively, such marker protein binding partners may also be found to be inhibitors of the marker protein.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that encodes a marker protein fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.,* GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo,* forming a marker-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.,* LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be readily detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the marker protein.

In a further embodiment, assays may be devised through the use of the invention for the purpose of identifying compounds which modulate (*e.g.,* affect either positively or negatively) interactions between a marker protein and its substrates and/or binding partners. Such compounds can include, but are not limited to, molecules such as antibodies, peptides, hormones, oligonucleotides, nucleic acids, and analogs thereof.

Such compounds may also be obtained from any available source, including systematic libraries of natural and/or synthetic compounds. The preferred assay components for use in this embodiment is an ovarian cancer marker protein identified herein, the known binding partner and/or substrate of same, and the test compound. Test compounds can be supplied from any source.

The basic principle of the assay systems used to identify compounds that interfere with the interaction between the marker protein and its binding partner involves preparing a reaction mixture containing the marker protein and its binding partner under conditions and for a time sufficient to allow the two products to interact and bind, thus forming a complex. In order to test an agent for inhibitory activity, the reaction mixture is prepared in the presence and absence of the test compound. The test compound can be initially included in the reaction mixture, or can be added at a time subsequent to the addition of the marker protein and its binding partner. Control reaction mixtures are incubated without the test compound or with a placebo. The formation of any complexes between the marker protein and its binding partner is then detected. The formation of a complex in the control reaction, but less or no such formation in the reaction mixture containing the test compound, indicates that the compound interferes with the interaction of the marker protein and its binding partner. Conversely, the formation of more complex in the presence of compound than in the control reaction indicates that the compound may enhance interaction of the marker protein and its binding partner.

The assay for compounds that interfere with the interaction of the marker protien with its binding partner may be conducted in a heterogeneous or homogeneous format. Heterogeneous assays involve anchoring either the marker protein or its binding partner onto a solid phase and detecting complexes anchored to the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the interaction between the marker proteins and the binding partners (*e.g.,* by competition) can be identified by conducting the reaction in the presence of the test substance, *i.e.,* by adding the test substance to the reaction mixture prior to or simultaneously with the marker and its interactive binding partner. Alternatively, test

compounds that disrupt preformed complexes, *e.g.*, compounds with higher binding constants that displace one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed. The various formats are briefly described below.

In a heterogeneous assay system, either the marker protein or its binding partner is anchored onto a solid surface or matrix, while the other corresponding non-anchored component may be labeled, either directly or indirectly. In practice, microtitre plates are often utilized for this approach. The anchored species can be immobilized by a number of methods, either non-covalent or covalent, that are typically well known to one who practices the art. Non-covalent attachment can often be accomplished simply by coating the solid surface with a solution of the marker protein or its binding partner and drying. Alternatively, an immobilized antibody specific for the assay component to be anchored can be used for this purpose. Such surfaces can often be prepared in advance and stored.

In related embodiments, a fusion protein can be provided which adds a domain that allows one or both of the assay components to be anchored to a matrix. For example, glutathione-S-transferase/marker fusion proteins or glutathione-S-transferase/binding partner can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and either the non-adsorbed marker or its binding partner, and the mixture incubated under conditions conducive to complex formation (*e.g.*, physiological conditions). Following incubation, the beads or microtiter plate wells are washed to remove any unbound assay components, the immobilized complex assessed either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of marker binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a marker protein or a marker protein binding partner can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated marker protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of

streptavidin-coated 96 well plates (Pierce Chemical). In certain embodiments, the protein-immobilized surfaces can be prepared in advance and stored.

In order to conduct the assay, the corresponding partner of the immobilized assay component is exposed to the coated surface with or without the test compound. After the reaction is complete, unreacted assay components are removed (*e.g.*, by washing) and any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the non-immobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the non-immobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; *e.g.*, using a labeled antibody specific for the initially non-immobilized species (the antibody, in turn, can be directly labeled or indirectly labeled with, *e.g.*, a labeled anti-Ig antibody). Depending upon the order of addition of reaction components, test compounds which modulate (inhibit or enhance) complex formation or which disrupt preformed complexes can be detected.

In an alternate embodiment of the invention, a homogeneous assay may be used. This is typically a reaction, analogous to those mentioned above, which is conducted in a liquid phase in the presence or absence of the test compound. The formed complexes are then separated from unreacted components, and the amount of complex formed is determined. As mentioned for heterogeneous assay systems, the order of addition of reactants to the liquid phase can yield information about which test compounds modulate (inhibit or enhance) complex formation and which disrupt preformed complexes.

In such a homogeneous assay, the reaction products may be separated from unreacted assay components by any of a number of standard techniques, including but not limited to: differential centrifugation, chromatography, electrophoresis and immunoprecipitation. In differential centrifugation, complexes of molecules may be separated from uncomplexed molecules through a series of centrifugal steps, due to the different sedimentation equilibria of complexes based on their different sizes and densities (see, for example, Rivas, G., and Minton, A.P., *Trends Biochem Sci* 1993 Aug;18(8):284-7). Standard chromatographic techniques may also be utilized to separate complexed molecules from uncomplexed ones. For example, gel filtration

chromatography separates molecules based on size, and through the utilization of an appropriate gel filtration resin in a column format, for example, the relatively larger complex may be separated from the relatively smaller uncomplexed components. Similarly, the relatively different charge properties of the complex as compared to the uncomplexed molecules may be exploited to differentially separate the complex from the remaining individual reactants, for example through the use of ion-exchange chromatography resins. Such resins and chromatographic techniques are well known to one skilled in the art (see, *e.g.,* Heegaard, 1998, *J Mol. Recognit.* 11:141-148; Hage and Tweed, 1997, *J. Chromatogr. B. Biomed. Sci. Appl.,* 699:499-525). Gel electrophoresis may also be employed to separate complexed molecules from unbound species (see, *e.g.,* Ausubel *et al* (eds.), as described in : Current Protocols in Molecular Biology, J. Wiley & Sons, New York. 1999). In this technique, protein or nucleic acid complexes are separated based on size or charge, for example. In order to maintain the binding interaction during the electrophoretic process, nondenaturing gels in the absence of reducing agent are typically preferred, but conditions appropriate to the particular interactants will be well known to one skilled in the art. Immunoprecipitation is another common technique utilized for the isolation of a protein-protein complex from solution (see, *e.g.,* Ausubel *et al* (eds.), In: Current Protocols in Molecular Biology, J. Wiley & Sons, New York. 1999). In this technique, all proteins binding to an antibody specific to one of the binding molecules are precipitated from solution by conjugating the antibody to a polymer bead that may be readily collected by centrifugation. The bound assay components are released from the beads (through a specific proteolysis event or other technique well known in the art which will not disturb the protein-protein interaction in the complex), and a second immunoprecipitation step is performed, this time utilizing antibodies specific for the correspondingly different interacting assay component. In this manner, only formed complexes should remain attached to the beads. Variations in complex formation in both the presence and the absence of a test compound can be compared, thus offering information about the ability of the compound to modulate interactions between the marker protein and its binding partner.

Also within the scope of the present invention are methods for direct detection of interactions between the marker protein and its natural binding partner and/or a test compound in a homogeneous or heterogeneous assay system without

further sample manipulation. For example, the technique of fluorescence energy transfer may be utilized (see, *e.g.,* Lakowicz *et al,* U.S. Patent No. 5,631,169; Stavrianopoulos *et al,* U.S. Patent No. 4,868,103). Generally, this technique involves the addition of a fluorophore label on a first 'donor' molecule (*e.g.,* marker or test compound) such that its emitted fluorescent energy will be absorbed by a fluorescent label on a second, 'acceptor' molecule (*e.g.,* marker or test compound), which in turn is able to fluoresce due to the absorbed energy. Alternately, the 'donor' protein molecule may simply utilize the natural fluorescent energy of tryptophan residues. Labels are chosen that emit different wavelengths of light, such that the 'acceptor' molecule label may be differentiated from that of the 'donor'. Since the efficiency of energy transfer between the labels is related to the distance separating the molecules, spatial relationships between the molecules can be assessed. In a situation in which binding occurs between the molecules, the fluorescent emission of the 'acceptor' molecule label in the assay should be maximal. An FET binding event can be conveniently measured through standard fluorometric detection means well known in the art (*e.g.,* using a fluorimeter). A test substance which either enhances or hinders participation of one of the species in the preformed complex will result in the generation of a signal variant to that of background. In this way, test substances that modulate interactions between a marker and its binding partner can be identified in controlled assays.

In another embodiment, modulators of marker expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of marker mRNA or protein in the cell, is determined. The level of expression of marker mRNA or protein in the presence of the candidate compound is compared to the level of expression of marker mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of marker expression based on this comparison. For example, when expression of marker mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of marker mRNA or protein expression. Conversely, when expression of marker mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of marker mRNA or protein expression. The level of marker mRNA or protein expression

in the cells can be determined by methods described herein for detecting marker mRNA or protein.

In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of a marker protein can be further confirmed *in vivo*, *e.g.*, in a whole animal model for cellular transformation and/or tumorigenesis.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, an marker modulating agent, an antisense marker nucleic acid molecule, an marker-specific antibody, or an marker-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

It is understood that appropriate doses of small molecule agents and protein or polypeptide agents depends upon a number of factors within the knowledge of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of these agents will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the agent to have upon the nucleic acid or polypeptide of the invention. Exemplary doses of a small molecule include milligram or microgram amounts per kilogram of subject or sample weight (*e.g.* about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram). Exemplary doses of a protein or polypeptide include gram, milligram or microgram amounts per kilogram of subject or sample weight (*e.g.* about 1 microgram per kilogram to about 5 grams per kilogram, about 100 micrograms per kilogram to about 500 milligrams per kilogram, or about 1 milligram per kilogram to about 50 milligrams per kilogram). It is furthermore

understood that appropriate doses of one of these agents depend upon the potency of the agent with respect to the expression or activity to be modulated. Such appropriate doses can be determined using the assays described herein. When one or more of these agents is to be administered to an animal (*e.g.* a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher can, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific agent employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediamine-tetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL (BASF; Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy

syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, or sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a polypeptide or antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium, and then incorporating the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed.

Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches, and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from a pressurized container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes having monoclonal antibodies incorporated therein or thereon) can also be used as pharmaceutically

acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

For antibodies, the preferred dosage is 0.1 mg/kg to 100 mg/kg of body weight (generally 10 mg/kg to 20 mg/kg). If the antibody is to act in the brain, a dosage of 50 mg/kg to 100 mg/kg is usually appropriate. Generally, partially human antibodies and fully human antibodies have a longer half-life within the human body than other antibodies. Accordingly, lower dosages and less frequent administration is often possible. Modifications such as lipidation can be used to stabilize antibodies and to enhance uptake and tissue penetration (*e.g.*, into the ovarian epithelium). A method for lipidation of antibodies is described by Cruikshank *et al.* (1997) *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14:193.

The invention also provides vaccine compositions for the prevention and/or treatment of ovarian cancer. The invention provides ovarian cancer vaccine compositions in which a protein of a marker of Table 1, or a combination of proteins of the markers of Table 1, are introduced into a subject in order to stimulate an immune response against the ovarian cancer. The invention also provides ovarian cancer vaccine compositions in which a gene expression construct, which expresses a marker or fragment of a marker identified in Table 1, is introduced into the subject such that a protein or fragment of a protein encoded by a marker of Table 1 is produced by transfected cells in the subject at a higher than normal level and elicits an immune response.

In one embodiment, an ovarian cancer vaccine is provided and employed as an immunotherapeutic agent for the prevention of ovarian cancer. In another embodiment, an ovarian cancer vaccine is provided and employed as an immunotherapeutic agent for the treatment of ovarian cancer.

By way of example, an ovarian cancer vaccine comprised of the proteins of the markers of Table 1, may be employed for the prevention and/or treatment of ovarian cancer in a subject by administering the vaccine by a variety of routes, *e.g.*, intradermally, subcutaneously, or intramuscularly. In addition, the ovarian cancer vaccine can be administered together with adjuvants and/or immunomodulators to boost the activity of the vaccine and the subject's response. In one embodiment, devices and/or compositions containing the vaccine, suitable for sustained or intermittent release could be, implanted in the body or topically applied thereto for the relatively slow release of such materials into the body. The ovarian cancer vaccine can be introduced along with immunomodulatory compounds, which can alter the type of immune response produced in order to produce a response which will be more effective in eliminating the cancer.

In another embodiment, an ovarian cancer vaccine comprised of an expression construct of the markers of Table 1, may be introduced by injection into muscle or by coating onto microprojectiles and using a device designed for the purpose to fire the projectiles at high speed into the skin. The cells of the subject will then express the protein(s) or fragments of proteins of the markers of Table 1 and induce an immune response. In addition, the ovarian cancer vaccine may be introduced along with expression constructs for immunomodulatory molecules, such as cytokines, which may increase the immune response or modulate the type of immune response produced in order to produce a response which will be more effective in eliminating the cancer.

The marker nucleic acid molecules of the present invention can also be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (U.S. Patent 5,328,470), or by stereotactic injection (see, *e.g.*, Chen *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively,

where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.* retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

V. Predictive Medicine

The present invention pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trails are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining the level of expression of one or more marker proteins or nucleic acids, in order to determine whether an individual is at risk of developing ovarian cancer. Such assays can be used for prognostic or predictive purposes to thereby prophylactically treat an individual prior to the onset of the cancer.

Yet another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs or other compounds administered either to inhibit ovarian cancer or to treat or prevent any other disorder {*i.e.* in order to understand any ovarian carcinogenic effects that such treatment may have}) on the expression or activity of a marker of the invention in clinical trials. These and other agents are described in further detail in the following sections.

A. Diagnostic Assays

An exemplary method for detecting the presence or absence of a marker protein or nucleic acid in a biological sample involves obtaining a biological sample (*e.g.* an ovary-associated body fluid) from a test subject and contacting the biological sample with a compound or an agent capable of detecting the polypeptide or nucleic acid (*e.g.*, mRNA, genomic DNA, or cDNA). The detection methods of the invention can thus be used to detect mRNA, protein, cDNA, or genomic DNA, for example, in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of a marker protein include enzyme linked immunosorbent

assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of a marker protein include introducing into a subject a labeled antibody directed against the protein or fragment thereof. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

A general principle of such diagnostic and prognostic assays involves preparing a sample or reaction mixture that may contain a marker, and a probe, under appropriate conditions and for a time sufficient to allow the marker and probe to interact and bind, thus forming a complex that can be removed and/or detected in the reaction mixture. These assays can be conducted in a variety of ways.

For example, one method to conduct such an assay would involve anchoring the marker or probe onto a solid phase support, also referred to as a substrate, and detecting target marker/probe complexes anchored on the solid phase at the end of the reaction. In one embodiment of such a method, a sample from a subject, which is to be assayed for presence and/or concentration of marker, can be anchored onto a carrier or solid phase support. In another embodiment, the reverse situation is possible, in which the probe can be anchored to a solid phase and a sample from a subject can be allowed to react as an unanchored component of the assay.

There are many established methods for anchoring assay components to a solid phase. These include, without limitation, marker or probe molecules which are immobilized through conjugation of biotin and streptavidin. Such biotinylated assay components can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). In certain embodiments, the surfaces with immobilized assay components can be prepared in advance and stored.

Other suitable carriers or solid phase supports for such assays include any material capable of binding the class of molecule to which the marker or probe belongs. Well-known supports or carriers include, but are not limited to, glass, polystyrene, nylon, polypropylene, nylon, polyethylene, dextran, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite.

In order to conduct assays with the above mentioned approaches, the non-immobilized component is added to the solid phase upon which the second component is anchored. After the reaction is complete, uncomplexed components may be removed (*e.g.*, by washing) under conditions such that any complexes formed will remain immobilized upon the solid phase. The detection of marker/probe complexes anchored to the solid phase can be accomplished in a number of methods outlined herein.

In a preferred embodiment, the probe, when it is the unanchored assay component, can be labeled for the purpose of detection and readout of the assay, either directly or indirectly, with detectable labels discussed herein and which are well-known to one skilled in the art.

It is also possible to directly detect marker/probe complex formation without further manipulation or labeling of either component (marker or probe), for example by utilizing the technique of fluorescence energy transfer (see, for example, Lakowicz *et al.*, U.S. Patent No. 5,631,169; Stavrianopoulos, *et al.*, U.S. Patent No. 4,868,103). A fluorophore label on the first, 'donor' molecule is selected such that, upon excitation with incident light of appropriate wavelength, its emitted fluorescent energy will be absorbed by a fluorescent label on a second 'acceptor' molecule, which in turn is able to fluoresce due to the absorbed energy. Alternately, the 'donor' protein molecule may simply utilize the natural fluorescent energy of tryptophan residues. Labels are chosen that emit different wavelengths of light, such that the 'acceptor' molecule label may be differentiated from that of the 'donor'. Since the efficiency of energy transfer between the labels is related to the distance separating the molecules, spatial relationships between the molecules can be assessed. In a situation in which binding occurs between the molecules, the fluorescent emission of the 'acceptor' molecule label in the assay should be maximal. An FET binding event can be conveniently measured through standard fluorometric detection means well known in the art (*e.g.*, using a fluorimeter).

In another embodiment, determination of the ability of a probe to recognize a marker can be accomplished without labeling either assay component (probe or marker) by utilizing a technology such as real-time Biomolecular Interaction Analysis (BIA) (see, *e.g.*, Sjolander, S. and Urbaniczky, C., 1991, *Anal. Chem.* 63:2338-2345

and Szabo *et al.*, 1995, *Curr. Opin. Struct. Biol.* 5:699-705). As used herein, "BIA" or "surface plasmon resonance" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the mass at the binding surface (indicative of a binding event) result in alterations of the refractive index of light near the surface (the optical phenomenon of surface plasmon resonance (SPR)), resulting in a detectable signal which can be used as an indication of real-time reactions between biological molecules.

Alternatively, in another embodiment, analogous diagnostic and prognostic assays can be conducted with marker and probe as solutes in a liquid phase. In such an assay, the complexed marker and probe are separated from uncomplexed components by any of a number of standard techniques, including but not limited to: differential centrifugation, chromatography, electrophoresis and immunoprecipitation. In differential centrifugation, marker/probe complexes may be separated from uncomplexed assay components through a series of centrifugal steps, due to the different sedimentation equilibria of complexes based on their different sizes and densities (see, for example, Rivas, G., and Minton, A.P., 1993, *Trends Biochem Sci.* 18(8):284-7). Standard chromatographic techniques may also be utilized to separate complexed molecules from uncomplexed ones. For example, gel filtration chromatography separates molecules based on size, and through the utilization of an appropriate gel filtration resin in a column format, for example, the relatively larger complex may be separated from the relatively smaller uncomplexed components. Similarly, the relatively different charge properties of the marker/probe complex as compared to the uncomplexed components may be exploited to differentiate the complex from uncomplexed components, for example through the utilization of ion-exchange chromatography resins. Such resins and chromatographic techniques are well known to one skilled in the art (see, *e.g.*, Heegaard, N.H., 1998, *J. Mol. Recognit.* Winter 11(1-6):141-8; Hage, D.S., and Tweed, S.A. *J Chromatogr B Biomed Sci Appl* 1997 Oct 10;699(1-2):499-525). Gel electrophoresis may also be employed to separate complexed assay components from unbound components (see, *e.g.*, Ausubel *et al.*, ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1987-1999). In this technique, protein or nucleic acid complexes are separated based on size or charge, for example. In order to maintain the binding interaction during the

electrophoretic process, non-denaturing gel matrix materials and conditions in the absence of reducing agent are typically preferred. Appropriate conditions to the particular assay and components thereof will be well known to one skilled in the art.

In a particular embodiment, the level of marker mRNA can be determined both by *in situ* and by *in vitro* formats in a biological sample using methods known in the art. The term "biological sample" is intended to include tissues, cells, biological fluids and isolates thereof, isolated from a subject, as well as tissues, cells and fluids present within a subject. Many expression detection methods use isolated RNA. For *in vitro* methods, any RNA isolation technique that does not select against the isolation of mRNA can be utilized for the purification of RNA from ovarian cells (see, *e.g.*, Ausubel *et al.*, ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York 1987-1999). Additionally, large numbers of tissue samples can readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski (1989, U.S. Patent No. 4,843,155).

The isolated mRNA can be used in hybridization or amplification assays that include, but are not limited to, Southern or Northern analyses, polymerase chain reaction analyses and probe arrays. One preferred diagnostic method for the detection of mRNA levels involves contacting the isolated mRNA with a nucleic acid molecule (probe) that can hybridize to the mRNA encoded by the gene being detected. The nucleic acid probe can be, for example, a full-length cDNA, or a portion thereof, such as an oligonucleotide of at least 7, 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to a mRNA or genomic DNA encoding a marker of the present invention. Other suitable probes for use in the diagnostic assays of the invention are described herein. Hybridization of an mRNA with the probe indicates that the marker in question is being expressed.

In one format, the mRNA is immobilized on a solid surface and contacted with a probe, for example by running the isolated mRNA on an agarose gel and transferring the mRNA from the gel to a membrane, such as nitrocellulose. In an alternative format, the probe(s) are immobilized on a solid surface and the mRNA is contacted with the probe(s), for example, in an Affymetrix gene chip array. A skilled

artisan can readily adapt known mRNA detection methods for use in detecting the level of mRNA encoded by the markers of the present invention.

An alternative method for determining the level of mRNA marker in a sample involves the process of nucleic acid amplification, *e.g.*, by rtPCR (the experimental embodiment set forth in Mullis, 1987, U.S. Patent No. 4,683,202), ligase chain reaction (Barany, 1991, *Proc. Natl. Acad. Sci. USA,* 88:189-193), self sustained sequence replication (Guatelli *et al.*, 1990, *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh *et al.*, 1989, *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi *et al.*, 1988, *Bio/Technology* 6:1197), rolling circle replication (Lizardi *et al.*, U.S. Patent No. 5,854,033) or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers. As used herein, amplification primers are defined as being a pair of nucleic acid molecules that can anneal to 5' or 3' regions of a gene (plus and minus strands, respectively, or vice-versa) and contain a short region in between. In general, amplification primers are from about 10 to 30 nucleotides in length and flank a region from about 50 to 200 nucleotides in length. Under appropriate conditions and with appropriate reagents, such primers permit the amplification of a nucleic acid molecule comprising the nucleotide sequence flanked by the primers.

For *in situ* methods, mRNA does not need to be isolated from the ovarian cells prior to detection. In such methods, a cell or tissue sample is prepared/processed using known histological methods. The sample is then immobilized on a support, typically a glass slide, and then contacted with a probe that can hybridize to mRNA that encodes the marker.

As an alternative to making determinations based on the absolute expression level of the marker, determinations may be based on the normalized expression level of the marker. Expression levels are normalized by correcting the absolute expression level of a marker by comparing its expression to the expression of a gene that is not a marker, *e.g.*, a housekeeping gene that is constitutively expressed. Suitable genes for normalization include housekeeping genes such as the actin gene, or epithelial cell-specific genes. This normalization allows the comparison of the

expression level in one sample, *e.g.*, a patient sample, to another sample, *e.g.*, a non-ovarian cancer sample, or between samples from different sources.

Alternatively, the expression level can be provided as a relative expression level. To determine a relative expression level of a marker, the level of expression of the marker is determined for 10 or more samples of normal versus cancer cell isolates, preferably 50 or more samples, prior to the determination of the expression level for the sample in question. The mean expression level of each of the genes assayed in the larger number of samples is determined and this is used as a baseline expression level for the marker. The expression level of the marker determined for the test sample (absolute level of expression) is then divided by the mean expression value obtained for that marker. This provides a relative expression level.

Preferably, the samples used in the baseline determination will be from ovarian cancer or from non-ovarian cancer cells of ovarian tissue. The choice of the cell source is dependent on the use of the relative expression level. Using expression found in normal tissues as a mean expression score aids in validating whether the marker assayed is ovarian specific (versus normal cells). In addition, as more data is accumulated, the mean expression value can be revised, providing improved relative expression values based on accumulated data. Expression data from ovarian cells provides a means for grading the severity of the ovarian cancer state.

In another embodiment of the present invention, a marker protein is detected. A preferred agent for detecting marker protein of the invention is an antibody capable of binding to such a protein or a fragment thereof, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment or derivatives thereof (*e.g.*, Fab or $F(ab')_2$) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin.

Proteins from ovarian cells can be isolated using techniques that are well known to those of skill in the art. The protein isolation methods employed can, for example, be such as those described in Harlow and Lane (Harlow and Lane, 1988, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York).

A variety of formats can be employed to determine whether a sample contains a protein that binds to a given antibody. Examples of such formats include, but are not limited to, enzyme immunoassay (EIA), radioimmunoassay (RIA), Western blot analysis and enzyme linked immunoabsorbant assay (ELISA). A skilled artisan can readily adapt known protein/antibody detection methods for use in determining whether ovarian cells express a marker of the present invention.

In one format, antibodies, or antibody fragments or derivatives, can be used in methods such as Western blots or immunofluorescence techniques to detect the expressed proteins. In such uses, it is generally preferable to immobilize either the antibody or proteins on a solid support. Suitable solid phase supports or carriers include any support capable of binding an antigen or an antibody. Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite.

One skilled in the art will know many other suitable carriers for binding antibody or antigen, and will be able to adapt such support for use with the present invention. For example, protein isolated from ovarian cells can be run on a polyacrylamide gel electrophoresis and immobilized onto a solid phase support such as nitrocellulose. The support can then be washed with suitable buffers followed by treatment with the detectably labeled antibody. The solid phase support can then be washed with the buffer a second time to remove unbound antibody. The amount of bound label on the solid support can then be detected by conventional means.

The invention also encompasses kits for detecting the presence of a marker protein or nucleic acid in a biological sample (*e.g.* an ovary-associated body fluid such as a urine sample). Such kits can be used to determine if a subject is suffering from or is at increased risk of developing ovarian cancer. For example, the kit can comprise a labeled compound or agent capable of detecting a marker protein or nucleic acid in a biological sample and means for determining the amount of the protein or

mRNA in the sample (*e.g.*, an antibody which binds the protein or a fragment thereof, or an oligonucleotide probe which binds to DNA or mRNA encoding the protein). Kits can also include instructions for interpreting the results obtained using the kit.

For antibody-based kits, the kit can comprise, for example: (1) a first antibody (*e.g.*, attached to a solid support) which binds to a marker protein; and, optionally, (2) a second, different antibody which binds to either the protein or the first antibody and is conjugated to a detectable label.

For oligonucleotide-based kits, the kit can comprise, for example: (1) an oligonucleotide, *e.g.*, a detectably labeled oligonucleotide, which hybridizes to a nucleic acid sequence encoding a marker protein or (2) a pair of primers useful for amplifying a marker nucleic acid molecule. The kit can also comprise, *e.g.*, a buffering agent, a preservative, or a protein stabilizing agent. The kit can further comprise components necessary for detecting the detectable label (*e.g.*, an enzyme or a substrate). The kit can also contain a control sample or a series of control samples which can be assayed and compared to the test sample. Each component of the kit can be enclosed within an individual container and all of the various containers can be within a single package, along with instructions for interpreting the results of the assays performed using the kit.

B. Pharmacogenomics

Agents or modulators which have a stimulatory or inhibitory effect on expression of a marker of the invention can be administered to individuals to treat (prophylactically or therapeutically) ovarian cancer in the patient. In conjunction with such treatment, the pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.*, drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the level of expression of a marker of the invention in an

individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, *e.g.*, Linder (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body are referred to as "altered drug action." Genetic conditions transmitted as single factors altering the way the body acts on drugs are referred to as "altered drug metabolism". These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is hemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, a PM will show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the level of expression of a marker of the invention in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a modulator of expression of a marker of the invention.

C. Monitoring Clinical Trials

Monitoring the influence of agents (*e.g.*, drug compounds) on the level of expression of a marker of the invention can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent to affect marker expression can be monitored in clinical trials of subjects receiving treatment for ovarian cancer. In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of one or more selected markers of the invention in the pre-administration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression of the marker(s) in the post-administration samples; (v) comparing the level of expression of the marker(s) in the pre-administration sample with the level of expression of the marker(s) in the post-administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent can be desirable to increase expression of the marker(s) to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent can be desirable to decrease expression of the marker(s) to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

D. Electronic Apparatus Readable Media and Arrays

Electronic apparatus readable media comprising a marker of the present invention is also provided. As used herein, "electronic apparatus readable media" refers to any suitable medium for storing, holding or containing data or information that can be read and accessed directly by an electronic apparatus. Such media can include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as compact disc; electronic storage media such as RAM, ROM, EPROM, EEPROM and the like; general hard disks and hybrids of these categories such as magnetic/optical storage media. The medium is adapted or configured for having recorded thereon a marker of the present invention.

As used herein, the term "electronic apparatus" is intended to include any suitable computing or processing apparatus or other device configured or adapted for storing data or information. Examples of electronic apparatus suitable for use with the present invention include stand-alone computing apparatus; networks, including a local area network (LAN), a wide area network (WAN) Internet, Intranet, and Extranet; electronic appliances such as a personal digital assistants (PDAs), cellular phone, pager and the like; and local and distributed processing systems.

As used herein, "recorded" refers to a process for storing or encoding information on the electronic apparatus readable medium. Those skilled in the art can readily adopt any of the presently known methods for recording information on known media to generate manufactures comprising the markers of the present invention.

A variety of software programs and formats can be used to store the marker information of the present invention on the electronic apparatus readable medium. For example, the marker nucleic acid sequence can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and MicroSoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like, as well as in other forms. Any number of data processor structuring formats (*e.g.*, text file or database) may be employed in order to obtain or create a medium having recorded thereon the the markers of the present invention.

By providing the markers of the invention in readable form, one can routinely access the marker sequence information for a variety of purposes. For example, one skilled in the art can use the nucleotide or amino acid sequences of the present invention in readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

The present invention therefore provides a medium for holding instructions for performing a method for determining whether a subject has ovarian cancer or a pre-disposition to ovarian cancer, wherein the method comprises the steps of determining the presence or absence of a marker and based on the presence or absence of the marker, determining whether the subject has ovarian cancer or a pre-disposition to ovarian cancer and/or recommending a particular treatment for ovarian cancer or pre-ovarian cancer condition.

The present invention further provides in an electronic system and/or in a network, a method for determining whether a subject has ovarian cancer or a pre-disposition to ovarian cancer associated with a marker wherein the method comprises the steps of determining the presence or absence of the marker, and based on the presence or absence of the marker, determining whether the subject has ovarian cancer or a pre-disposition to ovarian cancer, and/or recommending a particular treatment for the ovarian cancer or pre-ovarian cancer condition. The method may further comprise the step of receiving phenotypic information associated with the subject and/or acquiring from a network phenotypic information associated with the subject.

The present invention also provides in a network, a method for determining whether a subject has ovarian cancer or a pre-disposition to ovarian cancer associated with a marker, said method comprising the steps of receiving information associated with the marker receiving phenotypic information associated with the subject, acquiring information from the network corresponding to the marker and/or ovarian cancer, and based on one or more of the phenotypic information, the marker, and the acquired information, determining whether the subject has a ovarian cancer or a pre-disposition to ovarian cancer. The method may further comprise the step of

By providing the markers of the invention in readable form, one can routinely access the marker sequence information for a variety of purposes. For example, one skilled in the art can use the nucleotide or amino acid sequences of the present invention in readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

The present invention therefore provides a medium for holding instructions for performing a method for determining whether a subject has ovarian cancer or a pre-disposition to ovarian cancer, wherein the method comprises the steps of determining the presence or absence of a marker and based on the presence or absence of the marker, determining whether the subject has ovarian cancer or a pre-disposition to ovarian cancer and/or recommending a particular treatment for ovarian cancer or pre-ovarian cancer condition.

The present invention further provides in an electronic system and/or in a network, a method for determining whether a subject has ovarian cancer or a pre-disposition to ovarian cancer associated with a marker wherein the method comprises the steps of determining the presence or absence of the marker, and based on the presence or absence of the marker, determining whether the subject has ovarian cancer or a pre-disposition to ovarian cancer, and/or recommending a particular treatment for the ovarian cancer or pre-ovarian cancer condition. The method may further comprise the step of receiving phenotypic information associated with the subject and/or acquiring from a network phenotypic information associated with the subject.

The present invention also provides in a network, a method for determining whether a subject has ovarian cancer or a pre-disposition to ovarian cancer associated with a marker, said method comprising the steps of receiving information associated with the marker receiving phenotypic information associated with the subject, acquiring information from the network corresponding to the marker and/or ovarian cancer, and based on one or more of the phenotypic information, the marker, and the acquired information, determining whether the subject has a ovarian cancer or a pre-disposition to ovarian cancer. The method may further comprise the step of

recommending a particular treatment for the ovarian cancer or pre-ovarian cancer condition.

The present invention also provides a business method for determining whether a subject has ovarian cancer or a pre-disposition to ovarian cancer, said method comprising the steps of receiving information associated with the marker, receiving phenotypic information associated with the subject, acquiring information from the network corresponding to the marker and/or ovarian cancer, and based on one or more of the phenotypic information, the marker, and the acquired information, determining whether the subject has ovarian cancer or a pre-disposition to ovarian cancer. The method may further comprise the step of recommending a particular treatment for the ovarian cancer or pre-ovarian cancer condition.

The invention also includes an array comprising a marker of the present invention. The array can be used to assay expression of one or more genes in the array. In one embodiment, the array can be used to assay gene expression in a tissue to ascertain tissue specificity of genes in the array. In this manner, up to about 7600 genes can be simultaneously assayed for expression. This allows a profile to be developed showing a battery of genes specifically expressed in one or more tissues.

In addition to such qualitative determination, the invention allows the quantitation of gene expression. Thus, not only tissue specificity, but also the level of expression of a battery of genes in the tissue is ascertainable. Thus, genes can be grouped on the basis of their tissue expression *per se* and level of expression in that tissue. This is useful, for example, in ascertaining the relationship of gene expression between or among tissues. Thus, one tissue can be perturbed and the effect on gene expression in a second tissue can be determined. In this context, the effect of one cell type on another cell type in response to a biological stimulus can be determined. Such a determination is useful, for example, to know the effect of cell-cell interaction at the level of gene expression. If an agent is administered therapeutically to treat one cell type but has an undesirable effect on another cell type, the invention provides an assay to determine the molecular basis of the undesirable effect and thus provides the opportunity to co-administer a counteracting agent or otherwise treat the undesired effect. Similarly, even within a single cell type, undesirable biological effects can be

determined at the molecular level. Thus, the effects of an agent on expression of other than the target gene can be ascertained and counteracted.

In another embodiment, the array can be used to monitor the time course of expression of one or more genes in the array. This can occur in various biological contexts, as disclosed herein, for example development of ovarian cancer, progression of ovarian cancer, and processes, such a cellular transformation associated with ovarian cancer.

The array is also useful for ascertaining the effect of the expression of a gene on the expression of other genes in the same cell or in different cells. This provides, for example, for a selection of alternate molecular targets for therapeutic intervention if the ultimate or downstream target cannot be regulated.

The array is also useful for ascertaining differential expression patterns of one or more genes in normal and abnormal cells. This provides a battery of genes that could serve as a molecular target for diagnosis or therapeutic intervention.

E. Surrogate Markers

The markers of the invention may serve as surrogate markers for one or more disorders or disease states or for conditions leading up to disease states, and in particular, ovarian cancer. As used herein, a "surrogate marker" is an objective biochemical marker which correlates with the absence or presence of a disease or disorder, or with the progression of a disease or disorder (*e.g.*, with the presence or absence of a tumor). The presence or quantity of such markers is independent of the disease. Therefore, these markers may serve to indicate whether a particular course of treatment is effective in lessening a disease state or disorder. Surrogate markers are of particular use when the presence or extent of a disease state or disorder is difficult to assess through standard methodologies (*e.g.*, early stage tumors), or when an assessment of disease progression is desired before a potentially dangerous clinical endpoint is reached (*e.g.*, an assessment of cardiovascular disease may be made using cholesterol levels as a surrogate marker, and an analysis of HIV infection may be made using HIV RNA levels as a surrogate marker, well in advance of the undesirable clinical outcomes of myocardial infarction or fully-developed AIDS). Examples of the use of surrogate

markers in the art include: Koomen *et al.* (2000) *J. Mass. Spectrom.* 35: 258-264; and James (1994) *AIDS Treatment News Archive* 209.

The markers of the invention are also useful as pharmacodynamic markers. As used herein, a "pharmacodynamic marker" is an objective biochemical marker which correlates specifically with drug effects. The presence or quantity of a pharmacodynamic marker is not related to the disease state or disorder for which the drug is being administered; therefore, the presence or quantity of the marker is indicative of the presence or activity of the drug in a subject. For example, a pharmacodynamic marker may be indicative of the concentration of the drug in a biological tissue, in that the marker is either expressed or transcribed or not expressed or transcribed in that tissue in relationship to the level of the drug. In this fashion, the distribution or uptake of the drug may be monitored by the pharmacodynamic marker. Similarly, the presence or quantity of the pharmacodynamic marker may be related to the presence or quantity of the metabolic product of a drug, such that the presence or quantity of the marker is indicative of the relative breakdown rate of the drug *in vivo*. Pharmacodynamic markers are of particular use in increasing the sensitivity of detection of drug effects, particularly when the drug is administered in low doses. Since even a small amount of a drug may be sufficient to activate multiple rounds of marker transcription or expression, the amplified marker may be in a quantity which is more readily detectable than the drug itself. Also, the marker may be more easily detected due to the nature of the marker itself; for example, using the methods described herein, antibodies may be employed in an immune-based detection system for a protein marker, or marker-specific radiolabeled probes may be used to detect a mRNA marker. Furthermore, the use of a pharmacodynamic marker may offer mechanism-based prediction of risk due to drug treatment beyond the range of possible direct observations. Examples of the use of pharmacodynamic markers in the art include: Matsuda *et al.* US 6,033,862; Hattis *et al.* (1991) *Env. Health Perspect.* 90: 229-238; Schentag (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3: S21-S24; and Nicolau (1999) *Am, J. Health-Syst. Pharm.* 56 Suppl. 3: S16-S20.

VI. Experimental Protocol for all OV markers and M352 - M360

A. Identification of markers

The markers of the present invention were identified by transcriptional profiling using mRNA from 9 normal ovarian epithelia, 11 stage I/II ovarian cancer tumors and 25 stage III/IV tumors. Clones having expression at least two-fold higher in ovarian tumors as compared to their expression in non-ovarian tumor tissues in at least 4 tumor samples were selected to have their protein-encoding transcript sequences determined.

B. Identification of Markers and Assembly of Their Sequences

Clones which displayed an increase in expression in ovarian tumor samples over the corresponding average expression of non-tumor samples were used for further study. Briefly, BLAST analysis, against both public and proprietary sequence databases, of EST sequences known to be associated with each clone was performed, either directly or in the context of automatically, high-stringency assembled contiguous sequences. An identification of protein sequence corresponding to the clone was accomplished by obtaining one of the following:

a) a direct match between the protein sequence and at least one EST sequence in one of its 6 possible translations;

b) a direct match between the nucleotide sequence for the mRNA corresponding to the protein sequence and at least one EST sequence;

c) a match between the protein sequence and a contiguous assembly (contig) of the EST sequences with other available EST sequences in the databases in one of its 6 possible translations; or

d) a match between the nucleotide sequence for the mRNA corresponding to the protein sequence and a contiguous assembly of the EST sequences with other available EST sequences in the databases in one of its 6 possible translations.

C. Identification of Markers Having Newly-Identified Nucleotide and Amino Acid Sequences.

The markers of Table 2 include newly-identified amino acid sequences. These sequences were found to be novel based on one of the following criteria:

a) the protein sequence found within available public databases was incomplete or erroneous, leading to the construction of an additional completed/corrected protein sequence that is not found as such in the public domain;

b) based on nucleotide evidence, variants of the protein sequence were additionally constructed that are not found as such in the public domain; or

c) the contig for the EST sequences did not match any known protein, so that a novel protein sequence was derived from an open reading frame of the contig.

VII. Experimental Protocol for M68, M103, M138, M185, M312, M327-M328, M400, M430-M480, M559, M571-M573, M575-M576, M578-M583, M585-594, and M604-M617

A. Identification of Markers and Assembly of Their Sequences

The markers of the present invention were identified by transcription profiling using mRNA from 67 ovarian tumors of various histotypes and stage and 96 non-ovarian tumor tissues including normal ovarian epithelium, benign conditions, other normal tissues, and other abnormal tissues. Clones having expression at least three-fold higher in at least 10% of ovarian tumors, as compared to their expression in non-ovarian tumor tissue, were designated as ovarian cancer specific markers. These cDNA clones were selected to have their protein-encoding transcript sequences determined. Briefly, BLAST analysis, against both public and proprietary sequence databases, of EST sequences known to be associated with each clone was performed, either directly or in the context of automatically, high-stringency assembled contiguous sequences. An identification of protein sequence corresponding to the clone was accomplished by obtaining one of the following:

a) a direct match between the protein sequence and at least one EST sequence in one of its 6 possible translations;

b) a direct match between the nucleotide sequence for the mRNA corresponding to the protein sequence and at least one EST sequence;

c) a match between the protein sequence and a contiguous assembly (contig) of the EST sequences with other available EST sequences in the databases in one of its 6 possible translations; or

d) a match between the nucleotide sequence for the mRNA corresponding to the protein sequence and a contiguous assembly of the EST sequences with other available EST sequences in the databases in one of its 6 possible translations.

B. Identification of Markers Having Newly-Identified Amino Acid Sequences.

The markers of Table 2 include newly-identified amino acid sequences. These sequences were found to be novel based on one of the following criteria:

a) the protein sequence found within available public databases was incomplete or erroneous, leading to the construction of an additional completed/corrected protein sequence that is not found as such in the public domain;

b) based on nucleotide evidence, variants of the protein sequence were additionally constructed that are not found as such in the public domain; or

c) the contig for the EST sequences did not match any known protein, so that a novel protein sequence was derived from an open reading frame of the contig.

VIII. Gene Expression Analysis

Total RNA from normal human tissue was obtained from commercial sources. The integrity of the RNA was verified by agarose gel electrophoresis and ethidium bromide staining. Cell lines were purchased from ATCC and grown under the conditions recommended by ATCC. Total RNA from a number of various cell lines was prepared using commercial kits (Qiagen). First strand cDNA was prepared using oligo-dT primer and standard conditions. Each RNA preparation was treated with DNase I (Ambion) at 37°C for 1 hour.

Novel gene expression was measured by TaqMan® quantitative PCR (Perkin Elmer Applied Biosystems) in cDNA prepared from the following normal human tissues: heart, kidney, skeletal muscle, pancreas, skin, dorsal root ganglion,

breast, ovary, prostate, salivary glands, lung, colon, liver and lymph node. Figure 1 graphically represents the results of the TaqMan® expression study. The columns labelled A to V depict the expression level observed for OV88 in the following tissues:

Column A: Heart, normal tissue
Column B: Heart, CHF tissue
Column C: Kidney, normal tissue
Column D: Skeletal muscle, normal tissue
Column E: Pancreas, normal tissue
Column F: Skin, normal tissue
Column G: Dorsal root, normal tissue
Column H: Breast, normal tissue
Column I: Breast, tumor tissue
Column J Ovary, normal tissue
Column K: Ovary, tumor tissue
Column L: Prostate, normal tissue
Column M: Prostate, tumor tissue
Column N: Salivary glands, normal tissue
Column O: Lung, normal tissue
Column P: Lung, tumor tissue
Column Q: Lung, COPD tissue
Column R: Colon, IBD tissue
Column S: Liver, normal tissue
Column T: Liver fibrosis
Column U: Lymph node, normal tissue
Column V: Positive control

IX. Summary of the Data Provided in the Tables

Tables 1-3 list the markers of the present invention. In the Tables the markers are identified with a name ("Marker"), the name the gene is commonly known by, if applicable ("Gene Name"), the Sequence Listing identifier of the cDNA sequence of a nucleotide transcript encoded by or corresponding to the marker ("SEQ ID NO (nts)"), the Sequence Listing identifier of the amino acid sequence of a protein encoded

by the nucleotide transcript ("SEQ ID NO (AAs)"), and the location of the protein coding sequence within the cDNA sequence ("CDS").

Table 1 lists all of the markers of the invention, which are over-expressed in ovarian cancer cells compared to normal (*i.e.*, non-cancerous) ovarian cells and comprises markers listed in Tables 2 and 3. Table 2 lists newly-identified nucleotide and amino acid sequences useful as ovarian cancer markers. Table 3 lists newly-identified nucleotide sequences useful as ovarian cancer markers.

Other Embodiments

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims:

What is claimed:

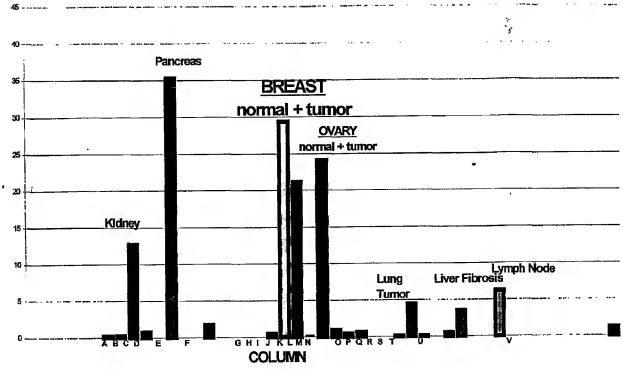
1. A method of assessing whether a patient is afflicted with ovarian cancer, the method comprising comparing:

a) the level of expression of a marker in a patient sample, wherein the marker is selected from Table 1, and

b) the normal level of expression of the marker in a control non-ovarian cancer sample,

wherein a significant increase in the level of expression of the marker in the patient sample and the normal level is an indication that the patient is afflicted with ovarian cancer.

Figure 1


45
40
35
30
25
20
15
10
5
0
Pancreas
BREAST
normal + tumor
OVARY
normal + tumor
Kidney
Lung
Tumor
Liver Fibrosis
Lymph Node
A B C D E F G H I J K L M N O P Q R S T U V
COLUMN

SEQUENCE LISTING

<110> Millennium Pharmaceuticals, Inc. et al.

<120> Nucleic Acid Molecules and Proteins For The Identification, Assessment, Prevention, and Therapy of Ovarian Cancer

<130> MRI-030PC

<150> 60/276,025
<151> 2001-03-14

<150> 60/325,149
<151> 2001-09-26

<150> 60/276,026
<151> 2001-03-14

<150> 60/324,967
<151> 2001/09/26

<150> 60/311,732
<151> 2001-08-10

<150> 60/325,102
<151> 2001-09-26

<150> 60/323,580
<151> 2001-09-19

<160> 363

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 4643
<212> DNA
<213> Homo sapiens

<400> 1
cctactctat tcagatattc tccagattcc taaagattag agatcatttc tcattctcct 60
aggagtactc acttcaggaa gcaaccagat aaaagagagg tgcaacggaa gccagaacat 120
tcctcctgga aattcaacct gtttcgcagt ttctcgagga atcagcattc agtcaatccg 180
ggccgggagc agtcatctgt ggtgaggctg attggctggg caggaacagc gccggggcgt 240
gggctgagca cagcgcttcg ctctctttgc cacaggaagc ctgagctcat tcgagtagcg 300
gctcttccaa gctcaaagaa gcagaggccg ctgttcgttt cctttaggtc tttccactaa 360
agtcggagta tcttcttcca agatttcacg tcttggtggc cgttccaagg agcgcgaggt 420
cgggatggat cttgaagggg accgcaatgg aggagcaaag aagaagaact tttttaaact 480
gaacaataaa agtgaaaaag ataagaagga aaagaaacca actgtcagtg tattttcaat 540
gtttcgctat tcaaattggc ttgacaagtt gtatatggtg gtgggaactt tggctgccat 600
catccatggg gctggacttc ctctcatgat gctggtgttt ggagaaatga cagatatctt 660
tgcaaatgca ggaaatttag aagatctgat gtcaaacatc actaatagaa gtgatatcaa 720
tgatacaggg ttcttcatga atctggagga agacatgacc aggtatgcct attattacag 780
tggaattggt gctggggtgc tggttgctgc ttacattcag gtttcatttt ggtgcctggc 840
agctggaaga caaatacaca aaattagaaa acagtttttt catgctataa tgcgacagga 900
gataggctgg tttgatgtgc acgatgttgg ggagcttaac acccgactta cagatgatgt 960
ctccaagatt aatgaaggaa ttggtgacaa aattggaatg ttctttcagt caatggcaac 1020

attttтcact gggtttatag taggatttac acgtggttgg aagctaaccc ttgtgatttt 1080
ggccatcagt cctgttcttg gactgtcagc tgctgtctgg gcaaagatac tatcttcatt 1140
tactgataaa gaactcttag cgtatgcaaa agctggagca gtagctgaag aggtcttggc 1200
agcaattaga actgtgattg catttggagg acaaaagaaa ggaacttgaa tgtgcaacaa 1260
aaatttagaa gaagctaaaa gaattgggat aaagaaagct attacagcca atatttctat 1320
aggtgctgct ttcctgctga tctatgcatc ttatgctctg gccttctggt atgcgaccac 1380
cttggtcctc tcaggggaat attctattgg acaagtactc actgtatttt ctgtattaat 1440
tggggctttt agtgttggac aggcatctcc aagcattgaa gcatttgcaa atgcaagagg 1500
agcagcttat gaaatcttca agataattga taataagcca agtattgaca gctattcgaa 1560
gagtgggcac aaaccagata atattaaggg aaatttggaa ttcagaaatg ttcacttcag 1620
ttacccatct cgaaaagaag ttaagatctt gaagggtctg aacctgaagg tgcagagtgg 1680
gcagacggtg gccctggttg gaaacagtgg ctgtgggaag agcacaacag tccagctgat 1740
gcagaggctc tatgacccca cagaggggat ggtcagtgtt gatggacagg atattaggac 1800
catаaatgta aggtttctac gggaaatcat tggtgtggtg agtcaggaac ctgtattgtt 1860
tgccaccacg atagctgaaa acattcgcta tggccgtgaa aatgtcacca tggatgagat 1920
tgagaaagct gtcaagyaag ccaatgccta tgacttatc atgaaactgc ctcataaatt 1980
tgacaccctg gttggagaga gaggggccca gttgagtggt gggcagaagc agaggatcgc 2040
cattgcacgt gccctggttc gcaaccccaa gatcctcctg ctggatgagg ccacgtcagc 2100
cttggacaca gaaagcgaag cagtggttca ggtggctctg gataaggcca gaaaaggtcy 2160
gaccaccatt gtgatagctc atcgtttgtc tacagttcgt aatgctgacg tcctgctgg 2220
tttcgatgat ggagtcattg tggagaaagg aaatcatgat gaactcatga aagagaaagg 2280
catttacttc aaacttgtca caatgcagac agcaggaaat gaagttgaat tagaaaatgc 2340
agctgatgaa tccaaaagtg aaattgatgc cttggaaatg tcttcaaatg attcaagatc 2400
cagtctaata agaaaaagat caactcgtag gagtgtccgt ggatcacaag cccaagacag 2460
aaagcttagt accaaagagg ctctggatga aagtatacct ccagtttcct tttggaggat 2520
tatgaagcta aatttaactg aatggcctta ttttgttgtt ggtgtatttt gtgccattat 2580
aaatggaggc ctgcaaccag catttgcaat aatattttca aagattatag gggtttttac 2640
aagaattgat gatcctgaaa caaaacgaca gaatagtaac ttgttttcac tattgtttct 2700
agcccttgga attatttctt ttattacatt tttccttcag ggtttcacat ttggcaaagc 2760
tggagagatc ctcaccaagc ggctccgata catggtttтc cgatccatgc tcagacagga 2820
tgtgagttgg tttgatgacc ctaaaaacac cactggagca ttgaclacca ggctcgccaa 2880
tgatgctgct caagttaaag gggctatagg ttccaggctt gctgtaatta cccagaatat 2940
agcaaatctt gggacaggaa taattatatc cttcatctat ggttggcaac taacactgtt 3000
actcttagca attgtaccca tcattgcaat agcaggagtt gttgaaatga aaatattgtc 3060
tggacaagca ctgaaagata agaaagaact agaaggtgct gggaagatcg ctactgaagc 3120
aatagaaaac ttccgaaccg ttgtttcttt gactcaggag cagaagtttg aacatatgta 3180
tgctcagagt ttgcaggtac catacagaaa ctctttgagg aaagcacaca tctttggaat 3240
tacatttcc ttcacccagg caatgatgta tttttcctat gctggatgtt tccggttgg 3300
agcctacttg gtggcacata aactcatgag ctttgaggat gttctgttag tattttcagc 3360
tgttgtcttt ggtgccatgg ccgtggggca agtcagttca tttgcctcctg actgtcataa 3420
agccaaaata tcagcagccc acatcatcat gatcattgaa aaaaccccтt tgattgacag 3480
ctacagcacg gaaggcctaa tgccgaacac attggaagga aatgtcacat ttggtgaagt 3540
tgtattcaac tatcccaccc gaccggacat cccagtgctt cagggactga gcctggaggt 3600
gaagaagggc cagacgctgg ctctggtggg cagcagtggc tgtggggaga gcacagtggt 3660
ccagctcctg gagcggttct acgacccctt ggcagggaaa gtgctgcttg atggcaaaga 3720
aataaagcga ctgaatgttc agtggctccg agcacacctg ggcatcgtgt cccaggagcc 3780
catcctgttt gactgcagca ttgctgagaa cattgcctat ggagacaaca gccgggtggt 3840
gtcacaggaa gagattgtga gggcagcaaa ggagggccaac atacatgcct tcatcgagtc 3900
actgcctaat aaatatagca ctaaagtagg agacaaagga actcagctct ctggtggcca 3960
gaaacaacgc attgccatag ctcgtgccct tgttagacag cctcatattt tgcttttgga 4020
tgaagccacg tcagctctgg atacagaaag tgaaaaggtt gtccaagaag ccctggacaa 4080
agccagagaa ggccgcacct gcattgtgat tgctcaccgc ctccagcaag aatgcc 4140
agacttaata gtggtgtttc agaatggcag agtcaaggag catggcacgc atcagcagct 4200
gctggcacag aaaggcatct atttttcaat ggtcagtgtc caggctggaa caaagcgcca 4260
gtgaactctg actgtatgag atgttaaata ctttttaata tttgtttaga tatgacattt 4320
attcaaagtt aaaagcaaac acttacagaa ttatgaagag gtatctgttt aacatttcct 4380
cagtcaagtt cagagtcttc agagacttcg taattaaagg aacagagtga gagacatcat 4440
caagtggaga gaaatcatag tttaaactgc attataaatt ttataacaga attaaagtag 4500

attttaaaag ataaaatgtg taattttgtt tatattttcc catttggact gtaactgact 4560
gccttgctaa aagattatag aagtagcaaa aagtattgaa atgtttgcat aaagtgtcta 4620
taataaaact aaactttcat gtg 4643

<210> 2
<211> 1279
<212> PRT
<213> Homo sapiens

<400> 2
Met Asp Leu Glu Gly Asp Arg Asn Gly Gly Ala Lys Lys Lys Asn Phe
1 5 10 15
Phe Lys Leu Asn Asn Lys Ser Glu Lys Asp Lys Lys Glu Lys Lys Pro
20 25 30
Thr Val Ser Val Phe Ser Met Phe Arg Tyr Ser Asn Trp Leu Asp Lys
35 40 45
Leu Tyr Met Val Val Gly Thr Leu Ala Ala Ile Ile His Gly Ala Gly
50 55 60
Leu Pro Leu Met Met Leu Val Phe Gly Glu Met Thr Asp Ile Phe Ala
65 70 75 80
Asn Ala Gly Asn Leu Glu Asp Leu Met Ser Asn Ile Thr Asn Arg Ser
85 90 95
Asp Ile Asn Asp Thr Gly Phe Phe Met Asn Leu Glu Glu Asp Met Thr
100 105 110
Arg Tyr Ala Tyr Tyr Tyr Ser Gly Ile Gly Ala Gly Val Leu Val Ala
115 120 125
Ala Tyr Ile Gln Val Ser Phe Trp Cys Leu Ala Ala Gly Arg Gln Ile
130 135 140
His Lys Ile Arg Lys Gln Phe Phe His Ala Ile Met Arg Gln Glu Ile
145 150 155 160
Gly Trp Phe Asp Val His Asp Val Gly Glu Leu Asn Thr Arg Leu Thr
165 170 175
Asp Asp Val Ser Lys Ile Asn Glu Gly Ile Gly Asp Lys Ile Gly Met
180 185 190
Phe Phe Gln Ser Met Ala Thr Phe Phe Thr Gly Phe Ile Val Gly Phe
195 200 205
Thr Arg Gly Trp Lys Leu Thr Leu Val Ile Leu Ala Ile Ser Pro Val
210 215 220
Leu Gly Leu Ser Ala Ala Val Trp Ala Lys Ile Leu Ser Ser Phe Thr
225 230 235 240
Asp Lys Glu Leu Leu Ala Tyr Ala Lys Ala Gly Ala Val Ala Glu Glu
245 250 255
Val Leu Ala Ala Ile Arg Thr Val Ile Ala Phe Gly Gly Gln Lys Lys
260 265 270
Glu Leu Glu Arg Tyr Asn Lys Asn Leu Glu Glu Ala Lys Arg Ile Gly
275 280 285
Ile Lys Lys Ala Ile Thr Ala Asn Ile Ser Ile Gly Ala Ala Phe Leu
290 295 300
Leu Ile Tyr Ala Ser Tyr Ala Leu Ala Phe Trp Tyr Gly Thr Thr Leu
305 310 315 320
Val Leu Ser Gly Glu Tyr Ser Ile Gly Gln Val Leu Thr Val Phe Ser
325 330 335
Val Leu Ile Gly Ala Phe Ser Val Gly Gln Ala Ser Pro Ser Ile Glu
340 345 350
Ala Phe Ala Asn Ala Arg Gly Ala Ala Tyr Glu Ile Phe Lys Ile Ile
355 360 365
Asp Asn Lys Pro Ser Ile Asp Ser Tyr Ser Lys Ser Gly His Lys Pro
370 375 380

```
Asp Asn Ile Lys Gly Asn Leu Glu Phe Arg Asn Val His Phe Ser Tyr
385                 390                 395                 400
Pro Ser Arg Lys Glu Val Lys Ile Leu Lys Gly Leu Asn Leu Lys Val
                405                 410                 415
Gln Ser Gly Gln Thr Val Ala Leu Val Gly Asn Ser Gly Cys Gly Lys
            420                 425                 430
Ser Thr Thr Val Gln Leu Met Gln Arg Leu Tyr Asp Pro Thr Glu Gly
        435                 440                 445
Met Val Ser Val Asp Gly Gln Asp Ile Arg Thr Ile Asn Val Arg Phe
    450                 455                 460
Leu Arg Glu Ile Ile Gly Val Val Ser Gln Glu Pro Val Leu Phe Ala
465                 470                 475                 480
Thr Thr Ile Ala Glu Asn Ile Arg Tyr Gly Arg Glu Asn Val Thr Met
                485                 490                 495
Asp Glu Ile Glu Lys Ala Val Lys Glu Ala Asn Ala Tyr Asp Phe Ile
            500                 505                 510
Met Lys Leu Pro His Lys Phe Asp Thr Leu Val Gly Glu Arg Gly Ala
        515                 520                 525
Gln Leu Ser Gly Gly Gln Lys Gln Arg Ile Ala Ile Ala Arg Ala Leu
    530                 535                 540
Val Arg Asn Pro Lys Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu
545                 550                 555                 560
Asp Thr Glu Ser Glu Ala Val Val Gln Val Ala Leu Asp Lys Ala Arg
                565                 570                 575
Lys Gly Arg Thr Thr Ile Val Ile Ala His Arg Leu Ser Thr Val Arg
            580                 585                 590
Asn Ala Asp Val Ile Ala Gly Phe Asp Asp Gly Val Ile Val Glu Lys
        595                 600                 605
Gly Asn His Asp Glu Leu Met Lys Glu Lys Gly Ile Tyr Phe Lys Leu
    610                 615                 620
Val Thr Met Gln Thr Ala Gly Asn Glu Val Glu Leu Glu Asn Ala Ala
625                 630                 635                 640
Asp Glu Ser Lys Ser Glu Ile Asp Ala Leu Glu Met Ser Ser Asn Asp
                645                 650                 655
Ser Arg Ser Ser Leu Ile Arg Lys Arg Ser Thr Arg Arg Ser Val Arg
            660                 665                 670
Gly Ser Gln Ala Gln Asp Arg Lys Leu Ser Thr Lys Glu Ala Leu Asp
        675                 680                 685
Glu Ser Ile Pro Pro Val Ser Phe Trp Arg Ile Met Lys Leu Asn Leu
    690                 695                 700
Thr Glu Trp Pro Tyr Phe Val Val Gly Val Phe Cys Ala Ile Ile Asn
705                 710                 715                 720
Gly Gly Leu Gln Pro Ala Phe Ala Ile Ile Phe Ser Lys Ile Ile Gly
                725                 730                 735
Val Phe Thr Arg Ile Asp Asp Pro Glu Thr Lys Arg Gln Asn Ser Asn
            740                 745                 750
Leu Phe Ser Leu Leu Phe Leu Ala Leu Gly Ile Ile Ser Phe Ile Thr
        755                 760                 765
Phe Phe Leu Gln Gly Phe Thr Phe Gly Lys Ala Gly Glu Ile Leu Thr
    770                 775                 780
Lys Arg Leu Arg Tyr Met Val Phe Arg Ser Met Leu Arg Gln Asp Val
785                 790                 795                 800
Ser Trp Phe Asp Asp Pro Lys Asn Thr Thr Gly Ala Leu Thr Thr Arg
                805                 810                 815
Leu Ala Asn Asp Ala Ala Gln Val Lys Gly Ala Ile Gly Ser Arg Leu
            820                 825                 830
Ala Val Ile Thr Gln Asn Ile Ala Asn Leu Gly Thr Gly Ile Ile Ile
        835                 840                 845
```

```
Ser Phe Ile Tyr Gly Trp Gln Leu Thr Leu Leu Leu Leu Ala Ile Val
    850                 855                 860
Pro Ile Ile Ala Ile Ala Gly Val Val Glu Met Lys Met Leu Ser Gly
865                 870                 875                 880
Gln Ala Leu Lys Asp Lys Lys Glu Leu Glu Gly Ala Gly Lys Ile Ala
                885                 890                 895
Thr Glu Ala Ile Glu Asn Phe Arg Thr Val Val Ser Leu Thr Gln Glu
            900                 905                 910
Gln Lys Phe Glu His Met Tyr Ala Gln Ser Leu Gln Val Pro Tyr Arg
        915                 920                 925
Asn Ser Leu Arg Lys Ala His Ile Phe Gly Ile Thr Phe Ser Phe Thr
    930                 935                 940
Gln Ala Met Met Tyr Phe Ser Tyr Ala Gly Cys Phe Arg Phe Gly Ala
945                 950                 955                 960
Tyr Leu Val Ala His Lys Leu Met Ser Phe Glu Asp Val Leu Leu Val
                965                 970                 975
Phe Ser Ala Val Val Phe Gly Ala Met Ala Val Gly Gln Val Ser Ser
            980                 985                 990
Phe Ala Pro Asp Tyr Ala Lys Ala Lys Ile Ser Ala Ala His Ile Ile
        995                 1000                1005
Met Ile Ile Glu Lys Thr Pro Leu Ile Asp Ser Tyr Ser Thr Glu Gly
    1010                1015                1020
Leu Met Pro Asn Thr Leu Glu Gly Asn Val Thr Phe Gly Glu Val Val
1025                1030                1035                1040
Phe Asn Tyr Pro Thr Arg Pro Asp Ile Pro Val Leu Gln Gly Leu Ser
                1045                1050                1055
Leu Glu Val Lys Lys Gly Gln Thr Leu Ala Leu Val Gly Ser Ser Gly
            1060                1065                1070
Cys Gly Lys Ser Thr Val Val Gln Leu Leu Glu Arg Phe Tyr Asp Pro
        1075                1080                1085
Leu Ala Gly Lys Val Leu Leu Asp Gly Lys Glu Ile Lys Arg Leu Asn
    1090                1095                1100
Val Gln Trp Leu Arg Ala His Leu Gly Ile Val Ser Gln Glu Pro Ile
1105                1110                1115                1120
Leu Phe Asp Cys Ser Ile Ala Glu Asn Ile Ala Tyr Gly Asp Asn Ser
                1125                1130                1135
Arg Val Val Ser Gln Glu Glu Ile Val Arg Ala Ala Lys Glu Ala Asn
            1140                1145                1150
Ile His Ala Phe Ile Glu Ser Leu Pro Asn Lys Tyr Ser Thr Lys Val
        1155                1160                1165
Gly Asp Lys Gly Thr Gln Leu Ser Gly Gly Gln Lys Gln Arg Ile Ala
    1170                1175                1180
Ile Ala Arg Ala Leu Val Arg Gln Pro His Ile Leu Leu Leu Asp Glu
1185                1190                1195                1200
Ala Thr Ser Ala Leu Asp Thr Glu Ser Glu Lys Val Val Gln Glu Ala
                1205                1210                1215
Leu Asp Lys Ala Arg Glu Gly Arg Thr Cys Ile Val Ile Ala His Arg
            1220                1225                1230
Leu Ser Thr Ile Gln Asn Ala Asp Leu Ile Val Val Phe Gln Asn Gly
        1235                1240                1245
Arg Val Lys Glu His Gly Thr His Gln Gln Leu Leu Ala Gln Lys Gly
    1250                1255                1260
Ile Tyr Phe Ser Met Val Ser Val Gln Ala Gly Thr Lys Arg Gln
1265                1270                1275
```

<210> 3
<211> 3859

<212> DNA
<213> Homo sapiens

<400> 3
aatctatcag ggaacggcgg tggccggtgc ggcgtgttcg gtgcgctctg gccgctcagg 60
ccgtgcggct gggtgagcgc acgcgaggcg gcgaggcggc atggtgtttt ctaggtcgtg 120
gcgtcgggct tccggagctt tggcggcagc tagggggaga tggcggagtc ttcggataag 180
ctctatcgag tcgagtacgc caagagcggg cgcgcctctt gcaagaaatg cagcgagagc 240
atccccaagg actcgctccg gatggccatc atggtgcagt cgcccatgtt tgatggaaaa 300
gtcccacact ggtaccactt ctcctgcttc tggaaggtgg gccactccat ccggcaccct 360
gacgttgagg tggatgggtt ctctgagctt cggtgggatg accagcagaa agtcaagaag 420
acagcggaag ctggaggagt gacaggcaaa ggccaggatg gaattggtag caaggcagag 480
aagactctgg gtgactttgc agcagagtat gccaagtcca acagaagtac gtgcaagggg 540
tgtatggaga agatagaaaa gggccaggtg cgcctgtcca agaagatggt ggacccggag 600
aagccacagc taggcatgat tgaccgctgg taccatccag gctgctttgt caagaacagg 660
gaggagctgg gtttccggcc cgagtacagt gcgagtcagc tcaaggcctt cagcctcctt 720
gctacagagg ataaagaagc cctgaagaag cagctcccag gagtcaagag tgaaggaaag 780
agaaaaggcg atgaggtgga tggagtggat gaagtggcga agaagaaatc taaaaaagaa 840
aaagacaagg atagtaagct tgaaaaagcc ctaaaggctc agaacgacct gatctggaac 900
atcaaggacg agctaaagaa agtgtgttca actaatgacc tgaaggagct actcatcttc 960
aacaagcagc aagtgccttc tggggagtcg gcgatcttgg accgagtagc tgatggcatg 1020
gtgttcggtg ccctccttcc ctgcgaggaa tgctcgggtc agctggtctt caagcagcgat 1080
gcctattact gcactgggga cgtcactgcc tggaccaagt gtatggtcaa gacacagaca 1140
cccaaccgga aggagtgggt aaccccaaag gaattccgag aaatctctta cctcaagaaa 1200
ttgaaggtta aaaagcagga ccgtatattc cccccagaaa ccagcgcctc cgtggcggcc 1260
acgcctccgc cctccacagc ctcggctcct gctgctgtga actcctctgc ttcagcagat 1320
aagccattat ccaacatgaa gatcctgact ctcgggaagc tgtcccggaa caaggatgaa 1380
gtgaaggcca tgattgagaa actcgggggg aagttgacgg ggacggccaa caaggcttcc 1440
ctgtgcatca gcaccaaaaa ggaggtggaa aagatgaata agaagatgga ggaagtaaag 1500
gaagccaaca tccgagttgt gtctgaggac ttcctccagg acgtctccgc ctccaccaag 1560
agccttcagg agttgttctt agcgcacatc ttgtcccctt gggaggcaga ggtgaaggca 1620
gagcctgttg aagttgtggc ccaagaggg aagtcaggag ctgcgctctc caaaaaaagc 1680
aagggccagg tcaaggagga aggtatcaac aaatctgaaa agagaatgaa attaactctt 1740
aaaggaggag cagctgtgga tcctgatttct ggactggaac actctgcgca tgtcctggag 1800
aaaggtggga aggtcttcag tgccaccctt ggcctggtgg acatcgttaa aggaaccaac 1860
tcctactaca agctgcagct tctggaggac gacaaggaaa acaggttatg gatattcagg 1920
tcctggggcc gtgtgggtac ggtgatcggt agcaacaaac tggaacagat gccgtccaag 1980
gaggatgcca tgtgagcagt catgaaatta tatgaagaaa aaaccgggaa cgttcggcac 2040
tccaaaaatt tcacgaagta tcccaaaaag ttttaccccc tggagattga ctatggccag 2100
gatgaagagg cagtgaagaa gctcacagta aatcctggca ccaagtccaa gctcccaag 2160
ccagttcagg acctcatcaa gatgatcttt gatgtggaaa gtatgaagaa agccatggtg 2220
gagtatgaga tgcaccttca gaagatgccc ttggggaagc tgagcaaaag gcagatccag 2280
gccgcatact ccatcctcag tgaggtccag caggcggtgt ctcagggcag cagcgacttct 2340
cagatcctgg atctctcaaa tcgcttttac accctgatcc cccacggctt tgggatgaag 2400
aagcctccgc tcctgaacaa tgcagacagt gtgcaggcca aggtggaaat gcttgacaac 2460
ctgctggaca tcgaggtggc ctacagtctg ctcagggggag ggtctgatga tagcagcaag 2520
gatcccatcg atgtcaacta tgagaagctc aaaactgaca ttaaggtggt tgacagagat 2580
tctgaagaag ccgagatcat caggaagtat gttaagaaca ctcatgcaac cacacacagt 2640
gcgtatgact tggaagtcat cgatatcttt aagatagagc gtgaaggcga atgccagcgt 2700
tacaagccct ttagcagct tcataaccga agattgctgt ggcacgggtc caggaccacc 2760
aactttgctg ggatcctgtc ccagggtctt cggatagccc cgcctgaagc gccgtgtaca 2820
ggctacatgt ttggtaaagg gatctatttc gctgacatgg tctccaagag tgccaactac 2880
taccatacgt ctcagggaga cccaataggc ttaatcctgt tgggagaagt tgcccttgga 2940
aacatgtatg aactgaagca cgcttcacat atcagcaggt taccccaggg caagcacagt 3000
gtcaaaggtt tgggcaaaac taccccctgat ccttcagcta acattagtct ggatggtgta 3060
gacgttcctc ttggagaccgg gatttcatct ggtgtgatag acaccctct actatataac 3120
gagtacattg tctatgatat tgctcaggta aatctgaagt atctgctgaa actgaaattc 3180
aattttaaga cctccctgtg gtaattggga gaggtagccg agtcacaccc ggtggctgtg 3240

gtatgaattc acccgaagcg cttctgcacc aactcacctg gccgctaagt tgctgatggg 3300
tagtacctgt actaaaccac ctcagaaagg attttacaga aacgtgttaa aggtttttctc 3360
taacttctca agtcccttgt tttgtgttgt gtctgtgggg aggggttgtt ttgggggttgt 3420
ttttgttttt tcttgccagg tagataaaac tgacatagag aaaaggctgg agagagattc 3480
tgttgcatag actagtccta tggaaaaaac caaagcttcg ttagaatgtc tgccttactg 3540
gtttccccag ggaaggaaaa atacacttcc acccttttttt ctaagtgttc gtctttagtt 3600
ttgattttgg aaagatgtta agcatttatt tttagttaaa ataaaaacta atttcatact 3660
atttagattt tctttttttat cttgcactta ttgtcccctt tttagttttt tttgtttgcc 3720
tcttgtggtg aggggtgtgg gaagaccaaa ggaaggaacg ctaacaattt ctcatactta 3780
gaaacaaaaa gagctttcct tctccaggaa tactgaacat gggagctctt gaaatatgta 3840
gtattaaaag ttgcatttg 3859

<210> 4
<211> 1014
<212> PRT
<213> Homo sapiens

<400> 4

Met	Ala	Glu	Ser	Ser	Asp	Lys	Leu	Tyr	Arg	Val	Glu	Tyr	Ala	Lys	Ser
1				5				10					15		
Gly	Arg	Ala	Ser	Cys	Lys	Lys	Cys	Ser	Glu	Ser	Ile	Pro	Lys	Asp	Ser
			20					25					30		
Leu	Arg	Met	Ala	Ile	Met	Val	Gln	Ser	Pro	Met	Phe	Asp	Gly	Lys	Val
		35					40					45			
Pro	His	Trp	Tyr	His	Phe	Ser	Cys	Phe	Trp	Lys	Val	Gly	His	Ser	Ile
	50					55					60				
Arg	His	Pro	Asp	Val	Glu	Val	Asp	Gly	Phe	Ser	Glu	Leu	Arg	Trp	Asp
65					70					75					80
Asp	Gln	Gln	Lys	Val	Lys	Lys	Thr	Ala	Glu	Ala	Gly	Gly	Val	Thr	Gly
				85					90					95	
Lys	Gly	Gln	Asp	Gly	Ile	Gly	Ser	Lys	Ala	Glu	Lys	Thr	Leu	Gly	Asp
			100					105					110		
Phe	Ala	Ala	Glu	Tyr	Ala	Lys	Ser	Asn	Arg	Ser	Thr	Cys	Lys	Gly	Cys
		115					120					125			
Met	Glu	Lys	Ile	Glu	Lys	Gly	Gln	Val	Arg	Leu	Ser	Lys	Lys	Met	Val
	130					135					140				
Asp	Pro	Glu	Lys	Pro	Gln	Leu	Gly	Met	Ile	Asp	Arg	Trp	Tyr	His	Pro
145					150					155					160
Gly	Cys	Phe	Val	Lys	Asn	Arg	Glu	Glu	Leu	Gly	Phe	Arg	Pro	Glu	Tyr
				165					170					175	
Ser	Ala	Ser	Gln	Leu	Lys	Gly	Phe	Ser	Leu	Leu	Ala	Thr	Glu	Asp	Lys
			180					185					190		
Glu	Ala	Leu	Lys	Lys	Gln	Leu	Pro	Gly	Val	Lys	Ser	Glu	Gly	Lys	Arg
		195					200					205			
Lys	Gly	Asp	Glu	Val	Asp	Gly	Val	Asp	Glu	Val	Ala	Lys	Lys	Lys	Ser
	210					215					220				
Lys	Lys	Glu	Lys	Asp	Lys	Asp	Ser	Lys	Leu	Glu	Lys	Ala	Leu	Lys	Ala
225					230					235					240
Gln	Asn	Asp	Leu	Ile	Trp	Asn	Ile	Lys	Asp	Glu	Leu	Lys	Lys	Val	Cys
				245					250					255	
Ser	Thr	Asn	Asp	Leu	Lys	Glu	Leu	Leu	Ile	Phe	Asn	Lys	Gln	Gln	Val
			260					265					270		
Pro	Ser	Gly	Glu	Ser	Ala	Ile	Leu	Asp	Arg	Val	Ala	Asp	Gly	Met	Val
		275					280					285			
Phe	Gly	Ala	Leu	Leu	Pro	Cys	Glu	Glu	Cys	Ser	Gly	Gln	Leu	Val	Phe
	290					295					300				
Lys	Ser	Asp	Ala	Tyr	Tyr	Cys	Thr	Gly	Asp	Val	Thr	Ala	Trp	Thr	Lys
305					310					315					320

```
Cys Met Val Lys Thr Gln Thr Pro Asn Arg Lys Glu Trp Val Thr Pro
                325                 330                 335
Lys Glu Phe Arg Glu Ile Ser Tyr Leu Lys Lys Leu Lys Val Lys Lys
            340                 345                 350
Gln Asp Arg Ile Phe Pro Pro Glu Thr Ser Ala Ser Val Ala Ala Thr
        355                 360                 365
Pro Pro Pro Ser Thr Ala Ser Ala Pro Ala Ala Val Asn Ser Ser Ala
    370                 375                 380
Ser Ala Asp Lys Pro Leu Ser Asn Met Lys Ile Leu Thr Leu Gly Lys
385                 390                 395                 400
Leu Ser Arg Asn Lys Asp Glu Val Lys Ala Met Ile Glu Lys Leu Gly
                405                 410                 415
Gly Lys Leu Thr Gly Thr Ala Asn Lys Ala Ser Leu Cys Ile Ser Thr
            420                 425                 430
Lys Lys Glu Val Glu Lys Met Asn Lys Lys Met Glu Glu Val Lys Glu
        435                 440                 445
Ala Asn Ile Arg Val Val Ser Glu Asp Phe Leu Gln Asp Val Ser Ala
    450                 455                 460
Ser Thr Lys Ser Leu Gln Glu Leu Phe Leu Ala His Ile Leu Ser Pro
465                 470                 475                 480
Trp Gly Ala Glu Val Lys Ala Glu Pro Val Glu Val Val Ala Pro Arg
                485                 490                 495
Gly Lys Ser Gly Ala Ala Leu Ser Lys Lys Ser Lys Gly Gln Val Lys
            500                 505                 510
Glu Glu Gly Ile Asn Lys Ser Glu Lys Arg Met Lys Leu Thr Leu Lys
        515                 520                 525
Gly Gly Ala Ala Val Asp Pro Asp Ser Gly Leu Glu His Ser Ala His
    530                 535                 540
Val Leu Glu Lys Gly Gly Lys Val Phe Ser Ala Thr Leu Gly Leu Val
545                 550                 555                 560
Asp Ile Val Lys Gly Thr Asn Ser Tyr Tyr Lys Leu Gln Leu Leu Glu
                565                 570                 575
Asp Asp Lys Glu Asn Arg Tyr Trp Ile Phe Arg Ser Trp Gly Arg Val
            580                 585                 590
Gly Thr Val Ile Gly Ser Asn Lys Leu Glu Gln Met Pro Ser Lys Glu
        595                 600                 605
Asp Ala Ile Glu Gln Phe Met Lys Leu Tyr Glu Glu Lys Thr Gly Asn
    610                 615                 620
Ala Trp His Ser Lys Asn Phe Thr Lys Tyr Pro Lys Lys Phe Tyr Pro
625                 630                 635                 640
Leu Glu Ile Asp Tyr Gly Gln Asp Glu Glu Ala Val Lys Lys Leu Thr
                645                 650                 655
Val Asn Pro Gly Thr Lys Ser Lys Leu Pro Lys Pro Val Gln Asp Leu
            660                 665                 670
Ile Lys Met Ile Phe Asp Val Glu Ser Met Lys Lys Ala Met Val Glu
        675                 680                 685
Tyr Glu Ile Asp Leu Gln Lys Met Pro Leu Gly Lys Leu Ser Lys Arg
    690                 695                 700
Gln Ile Gln Ala Ala Tyr Ser Ile Leu Ser Glu Val Gln Gln Ala Val
705                 710                 715                 720
Ser Gln Gly Ser Ser Asp Ser Gln Ile Leu Asp Leu Ser Asn Arg Phe
                725                 730                 735
Tyr Thr Leu Ile Pro His Asp Phe Gly Met Lys Lys Pro Pro Leu Leu
            740                 745                 750
Asn Asn Ala Asp Ser Val Gln Ala Lys Val Glu Met Leu Asp Asn Leu
        755                 760                 765
Leu Asp Ile Glu Val Ala Tyr Ser Leu Leu Arg Gly Gly Ser Asp Asp
    770                 775                 780
```

```
Ser Ser Lys Asp Pro Ile Asp Val Asn Tyr Glu Lys Leu Lys Thr Asp
785                 790                 795                 800
Ile Lys Val Val Asp Arg Asp Ser Glu Glu Ala Glu Ile Ile Arg Lys
                805                 810                 815
Tyr Val Lys Asn Thr His Ala Thr Thr His Ser Ala Tyr Asp Leu Glu
            820                 825                 830
Val Ile Asp Ile Phe Lys Ile Glu Arg Glu Gly Glu Cys Gln Arg Tyr
        835                 840                 845
Lys Pro Phe Lys Gln Leu His Asn Arg Arg Leu Leu Trp His Gly Ser
    850                 855                 860
Arg Thr Thr Asn Phe Ala Gly Ile Leu Ser Gln Gly Leu Arg Ile Ala
865                 870                 875                 880
Pro Pro Glu Ala Pro Val Thr Gly Tyr Met Phe Gly Lys Gly Ile Tyr
                885                 890                 895
Phe Ala Asp Met Val Ser Lys Ser Ala Asn Tyr Tyr His Thr Ser Gln
            900                 905                 910
Gly Asp Pro Ile Gly Leu Ile Leu Leu Gly Glu Val Ala Leu Gly Asn
        915                 920                 925
Met Tyr Glu Leu Lys His Ala Ser His Ile Ser Arg Leu Pro Lys Gly
    930                 935                 940
Lys His Ser Val Lys Gly Leu Gly Lys Thr Thr Pro Asp Pro Ser Ala
945                 950                 955                 960
Asn Ile Ser Leu Asp Gly Val Asp Val Pro Leu Gly Thr Gly Ile Ser
                965                 970                 975
Ser Gly Val Ile Asp Thr Ser Leu Leu Tyr Asn Glu Tyr Ile Val Tyr
            980                 985                 990
Asp Ile Ala Gln Val Asn Leu Lys Tyr Leu Leu Lys Leu Lys Phe Asn
        995                 1000                1005
Phe Lys Thr Ser Leu Trp
    1010
```

<210> 5
<211> 1465
<212> DNA
<213> Homo sapiens

<400> 5

```
ggcacgaggg ggacgctctc agctctcggc gcacggccca ggttatcttg tagcatagca 60
acttcggatt tcactctacc cggagagttt cccgcttggt tgaacacatt ggcctcagga 120
agcttccttc aaaatgtcta ctgttcacga aatcctgtgc aagctcagct tggagggtga 180
tcactctaca cccccaagtg catatgggtc tgtcaaagcc tatactaact ttgatgctga 240
gcgggatgct ttgaacattg aaacagccat caagaccaaa ggtgtggatg aggtcaccat 300
tgtcaacatt ttgaccaacc gcagcaatgc acagagacag gatattgcct tcgcctacca 360
gagaaggacc aaaaaggaac ttgcatcagc actgaagtca gccttatctg gccacctgga 420
gacggtgatt ttgggcctat tgaagacacc tgctcagtat gacgcttctg agctaaaagc 480
ttccatgaag gggctgggaa ccgacgagga ctctctcatt gagatcatct gctccagaac 540
caaccaggag ctgcaggaaa ttaacagagt ctacaaggaa atgtacaaga ctgatctgga 600
gaaggacatt atttcggaca catctggtga cttccgcaag ctgatggttg ccctggcaaa 660
gggtagaaga gcagaggatg gctctgtcat tgattatgaa ctgattgacc aagatgctcg 720
ggatctctat gacgctggag tgaagaggaa aggaactgat gttcccaagt ggatcagcat 780
catgaccgag cggagcgtgc cccacctcca gaaagtattt gataggtaca agagttacag 840
cccttatgac atgttggaaa gcatcaggaa agaggttaaa ggagacctgg aaaatgcttt 900
cctgaacctg gttcagtgca ttcagaacaa gcccctgtat tttgctgatc ggctgtatga 960
ctccatgaag ggcaagggga cgcgagataa ggtcctgatc agaatcatgg tctccgcgag 1020
tgaagtggac atgttgaaaa ttaggtctga attcaagaga aagtacggca agtcccttga 1080
ctattatatc cagcaagaca ctaagggcga ctaccagaaa gcgctgctgt acctgtgtgg 1140
tggagatgac tgaagcccga cacggcctga gcgtccagaa atggtgctca ccatgcttcc 1200
```

agctaacagg tctagaaaac cagcttgcga ataacagtcc ccgtggccat ccctgtgagg 1260
gtgacgttag cattaccccc aacctcattt tagttgccta agcattgcct ggccttcctg 1320
tctagtctct cctgtaagcc aaagaaatga acattccaag gagttggaag tgaagtctat 1380
gatgtgaaac actttgcctc ctgtgtactg tgtcataaac agatgaataa actgaatttg 1440
tactttaaaa aaaaaaaaa aaaaa 1465

<210> 6
<211> 339
<212> PRT
<213> Homo sapiens

<400> 6
Met Ser Thr Val His Glu Ile Leu Cys Lys Leu Ser Leu Glu Gly Asp
1 5 10 15
His Ser Thr Pro Pro Ser Ala Tyr Gly Ser Val Lys Ala Tyr Thr Asn
20 25 30
Phe Asp Ala Glu Arg Asp Ala Leu Asn Ile Glu Thr Ala Ile Lys Thr
35 40 45
Lys Gly Val Asp Glu Val Thr Ile Val Asn Ile Leu Thr Asn Arg Ser
50 55 60
Asn Ala Gln Arg Gln Asp Ile Ala Phe Ala Tyr Gln Arg Arg Thr Lys
65 70 75 80
Lys Glu Leu Ala Ser Ala Leu Lys Ser Ala Leu Ser Gly His Leu Glu
85 90 95
Thr Val Ile Leu Gly Leu Leu Lys Thr Pro Ala Gln Tyr Asp Ala Ser
100 105 110
Glu Leu Lys Ala Ser Met Lys Gly Leu Gly Thr Asp Glu Asp Ser Leu
115 120 125
Ile Glu Ile Ile Cys Ser Arg Thr Asn Gln Glu Leu Gln Glu Ile Asn
130 135 140
Arg Val Tyr Lys Glu Met Tyr Lys Thr Asp Leu Glu Lys Asp Ile Ile
145 150 155 160
Ser Asp Thr Ser Gly Asp Phe Arg Lys Leu Met Val Ala Leu Ala Lys
165 170 175
Gly Arg Arg Ala Glu Asp Gly Ser Val Ile Asp Tyr Glu Leu Ile Asp
180 185 190
Gln Asp Ala Arg Asp Leu Tyr Asp Ala Gly Val Lys Arg Lys Gly Thr
195 200 205
Asp Val Pro Lys Trp Ile Ser Ile Met Thr Glu Arg Ser Val Pro His
210 215 220
Leu Gln Lys Val Phe Asp Arg Tyr Lys Ser Tyr Ser Pro Tyr Asp Met
225 230 235 240
Leu Glu Ser Ile Arg Lys Glu Val Bys Gly Asp Leu Glu Asn Ala Phe
245 250 255
Leu Asn Leu Val Gln Cys Ile Gln Asn Lys Pro Leu Tyr Phe Ala Asp
260 265 270
Arg Leu Tyr Asp Ser Met Lys Gly Lys Gly Thr Arg Asp Lys Val Leu
275 280 285
Ile Arg Ile Met Val Ser Arg Ser Glu Val Asp Met Leu Lys Ile Arg
290 295 300
Ser Glu Phe Lys Arg Lys Tyr Gly Lys Ser Leu Tyr Tyr Tyr Ile Gln
305 310 315 320
Gln Asp Thr Lys Gly Asp Tyr Gln Lys Ala Leu Leu Tyr Leu Cys Gly
325 330 335
Gly Asp Asp

<210> 7
<211> 1362
<212> DNA
<213> Homo sapiens

<400> 7
catttgggga cgctctcagc tctcggcgca cggcccagct tccttcaaaa tgtctactgt 60
tcacgaaatc ctgtgcaagc tcagcttgga gggtgatcac tctacacccc caagtgcata 120
tgggtctgtc aaagcctata ctaactttga tgctgagcgg gatgctttga acattgaaac 180
agccatcaag accaaaggtg tggatgaggt caccattgtc aacatttttga ccaacccgcag 240
caatgcacag agacaggata ttgccttcgc ctaccagaga aggaccaaaa aggaacttgc 300
atcagcactg aagtcagcct tatctggcca cctggagacg gtgattttgg gcctattgaa 360
gacacctgct cagtatgacg cttctgagct aaaagcttcc atgaaggggc tgggaaccga 420
cgaggactct ctcattgaga tcatctgctc cagaaccaac caggagctgc aggaaattaa 480
cagagtctac aaggaaatgt acaagactga tctggagaag gacattattt cggacacatc 540
tggtgacttc cgcaagctga tggttgccct ggcaaaggat agaagagcag aggatggctc 600
tgtcattgat tatgaactga ttgaccaaga tgctcgggat ctctatgacg ctggagtgaa 660
gaggaaagga actgatgttc ccaagtggat cagcatcatg accgagcgga gcgtgcccca 720
cctccagaaa gtatttgata ggtacaagag ttacagccct tatgacatgt tggaaagcat 780
caggaaagag gttaaaggag acctggaaaa tgctttcctg aacctggttc agtgcattca 840
gaacaagccc ctgtattttg ctgatcggct gtatgactcc atgaagggca aggggacgcg 900
agataaggtc ctgatcagaa tcatggtctc ccgcagtgaa gtggacatgt tgaaaattag 960
gtctgaattc aagagaaagt acggcaagtc cctgtactat tatatccagc aagacactaa 1020
gggcgactac cagaaagcgc tgctgtacct gtgtggtgga gatgactgaa gcccgacacg 1080
gcctgagcgt ccagaaatgg tgctcaccat gcttccagct aacaggtcta gaaaaccagc 1140
ttgcgaataa cagtcccctgt ggccatccct gtgagggtga cgttagcatt acccccaacc 1200
tcattttagt tgcctaagca ttgcctggcc ttcctgtcta gtctctcctg taagccaaag 1260
aaatgaacat tccaaggagt tggaagtgaa gtctatgatg tgaaacactt tgcctcctgt 1320
gtactgtgtc ataaacagat gaataaactg aatttgtact tt 1362

<210> 8
<211> 339
<212> PRT
<213> Homo sapiens

<400> 8
Met Ser Thr Val His Glu Ile Leu Cys Lys Leu Ser Leu Glu Gly Asp
1 5 10 15
His Ser Thr Pro Pro Ser Ala Tyr Gly Ser Val Lys Ala Tyr Thr Asn
20 25 30
Phe Asp Ala Glu Arg Asp Ala Leu Asn Ile Glu Thr Ala Ile Lys Thr
35 40 45
Lys Gly Val Asp Glu Val Thr Ile Val Asn Ile Leu Thr Asn Arg Ser
50 55 60
Asn Ala Gln Arg Gln Asp Ile Ala Phe Ala Tyr Gln Arg Arg Thr Lys
65 70 75 80
Lys Glu Leu Ala Ser Ala Leu Lys Ser Ala Leu Ser Gly His Leu Glu
85 90 95
Thr Val Ile Leu Gly Leu Leu Lys Thr Pro Ala Gln Tyr Asp Ala Ser
100 105 110
Glu Leu Lys Ala Ser Met Lys Gly Leu Gly Thr Asp Glu Asp Ser Leu
115 120 125
Ile Glu Ile Ile Cys Ser Arg Thr Asn Gln Glu Leu Gln Glu Ile Asn
130 135 140
Arg Val Tyr Lys Glu Met Tyr Lys Thr Asp Leu Glu Lys Asp Ile Ile
145 150 155 160
Ser Asp Thr Ser Gly Asp Phe Arg Lys Leu Met Val Ala Leu Ala Lys
165 170 175

```
Gly Arg Arg Ala Glu Asp Gly Ser Val Ile Asp Tyr Glu Leu Ile Asp
            180                 185                 190
Gln Asp Ala Arg Asp Leu Tyr Asp Ala Gly Val Lys Arg Lys Gly Thr
        195                 200                 205
Asp Val Pro Lys Trp Ile Ser Ile Met Thr Glu Arg Ser Val Pro His
    210                 215                 220
Leu Gln Lys Val Phe Asp Arg Tyr Lys Ser Tyr Ser Pro Tyr Asp Met
225                 230                 235                 240
Leu Glu Ser Ile Arg Lys Glu Val Lys Gly Asp Leu Glu Asn Ala Phe
                245                 250                 255
Leu Asn Leu Val Gln Cys Ile Gln Asn Lys Pro Leu Tyr Phe Ala Asp
            260                 265                 270
Arg Leu Tyr Asp Ser Met Lys Gly Lys Gly Thr Arg Asp Lys Val Leu
        275                 280                 285
Ile Arg Ile Met Val Ser Arg Ser Glu Val Asp Met Leu Lys Ile Arg
    290                 295                 300
Ser Glu Phe Lys Arg Lys Tyr Gly Lys Ser Leu Tyr Tyr Tyr Ile Gln
305                 310                 315                 320
Gln Asp Thr Lys Gly Asp Tyr Gln Lys Ala Leu Leu Tyr Leu Cys Gly
                325                 330                 335
Gly Asp Asp
```

<210> 9
<211> 1982
<212> DNA
<213> Homo sapiens

<400> 9

```
gcagaggagg agcgcacgcc ggcctcgaag aacttctgct tgggtggctg aactctgatc 60
ttgacctaga gtcatggcca tggcaaccaa aggaggtact gtcaaagctg cttcaggatt 120
caatgccatg gaagatgccc agaccctgag gaaggccatg aaagggctcg gcaccgatga 180
agacgccatt attagcgtcc ttgcctaccg caacaccgcc cagcgccagg agatcaggac 240
agcctacaag agcaccatcg gcagggactt gatagacgac ctgaagtcag aactgagtgg 300
caacttcgag caggtgattg tggggatgat gacgcccacg gtgctgtatg acgtgcaaga 360
gctgcgaagg gccatgaagg gagccggcac tgatgagggc tgcctaattg agatcctggc 420
ctccggacc cctgaggaga tccggcgcat aagccaaacc tacaagcagc aatatggacg 480
gagccttgaa gatgacattc gctctgacac atcgttcatg ttccagcgag tgctggtgtc 540
tctgtcagct ggtgggaggg atgaaggaaa ttatctgtac gatgctctcg tggacacgga 600
tgcccaggac ctgtatgagg ctggagagaa gaaatggggg acagatgagg tgaaatttct 660
aactgttctc tgttcccgga accgaaatca cctgttgcat gtgttttgatg aatacaaaag 720
gatatcacag aaggatattg aacagagtat taaatctgaa acatctggta gctttgaaga 780
tgctctgctg gctatagtaa agtgcatgag gaacaaatct gcatattttg ctgaaaagct 840
ctataaatcg atgaagggct tgggcaccga tgataacacc ctcatcagag tgatggtttc 900
tcgagcagaa attgacatgt tggatatccg ggcacacttc aagagactct atggaaagtc 960
tctgtactcg ttcatcaagg gtgacacatc tggagactac aggaaagtac tgcttgttct 1020
ctgtggagga gatgattaaa ataaaaatcc cagaaggaca ggaggattct caacactttg 1080
aatttttttta acttcatttt tctcacactgc tattatcatt atctcagaat gcttatttcc 1140
aattaaaacg cctacagctg cctcctagaa tatagactgt ctgtattatt attcacctat 1200
aattagtcat tatgatgctt taaagctgta cttgcatttc aaagcttata agatataaat 1260
ggagatttta aagtagaaat aaatatgtat tccatgtttt taaaagatta ctttctactt 1320
tgtgtttcac agacattgaa tatattaaat tattccatat ttttttttca gtgaaaaatt 1380
ttttaaatgg aagactgttc taaaatcact tttttcccta atccaatttt tagagtggct 1440
agtagtttct tcatttgaaa ttgtaagcat ccggtcagta agaatgccca tccagttttc 1500
tatatttcat agtcaaagcc ttgaaagcat ctacaaatct ctttttttag gttttgtcca 1560
tagcatcagt tgatccttac taagtttttc atgggagact tccttcatca catcttatgt 1620
tgaaatcact ttctgtagtc aaagtatacc aaaaccaatt tatctgaact aaattctaaa 1680
```

gtatggttat acaaaccata tacatctggt taccaaacat aaatgctgaa cattccatat 1740
tattatagtt aatgtcttaa tccagcttgc aagtgaatgg aaaaaaaaat aagcttcaaa 1800
ctaggtattc tgggaatgat gtaatgctct gaatttagta tgatataaag aaaacttttt 1860
tgtgctaaaa atacttttta aaatcaattt tgttgattgt agtaatttct atttgcactg 1920
tgcctttcaa ctccagaaac attctaagat gtacttggat ttaattaaaa agttcacttt 1980
gt 1982

<210> 10
<211> 321
<212> PRT
<213> Homo sapiens

<400> 10
Met Ala Met Ala Thr Lys Gly Gly Thr Val Lys Ala Ala Ser Gly Phe
1 5 10 15
Asn Ala Met Glu Asp Ala Gln Thr Leu Arg Lys Ala Met Lys Gly Leu
20 25 30
Gly Thr Asp Glu Asp Ala Ile Ile Ser Val Leu Ala Tyr Arg Asn Thr
35 40 45
Ala Gln Arg Gln Glu Ile Arg Thr Ala Tyr Lys Ser Thr Ile Gly Arg
50 55 60
Asp Leu Ile Asp Asp Leu Lys Ser Glu Leu Ser Gly Asn Phe Glu Gln
65 70 75 80
Val Ile Val Gly Met Met Thr Pro Thr Val Leu Tyr Asp Val Gln Glu
85 90 95
Leu Arg Arg Ala Met Lys Gly Ala Gly Thr Asp Glu Gly Cys Leu Ile
100 105 110
Glu Ile Leu Ala Ser Arg Thr Pro Glu Glu Ile Arg Arg Ile Ser Gln
115 120 125
Thr Tyr Gln Gln Gln Tyr Gly Arg Ser Leu Glu Asp Asp Ile Arg Ser
130 135 140
Asp Thr Ser Phe Met Phe Gln Arg Val Leu Val Ser Leu Ser Ala Gly
145 150 155 160
Gly Arg Asp Glu Gly Asn Tyr Leu Asp Asp Ala Leu Val Arg Gln Asp
165 170 175
Ala Gln Asp Leu Tyr Glu Ala Gly Glu Lys Lys Trp Gly Thr Asp Glu
180 185 190
Val Lys Phe Leu Thr Val Leu Cys Ser Arg Asn Arg Asn His Leu Leu
195 200 205
His Val Phe Asp Glu Tyr Lys Arg Ile Ser Gln Lys Asp Ile Glu Gln
210 215 220
Ser Ile Lys Ser Glu Thr Ser Gly Ser Phe Glu Asp Ala Leu Leu Ala
225 230 235 240
Ile Val Lys Cys Met Arg Asn Lys Ser Ala Tyr Phe Ala Glu Lys Leu
245 250 255
Tyr Lys Ser Met Lys Gly Leu Gly Thr Asp Asp Asn Thr Leu Ile Arg
260 265 270
Val Met Val Ser Arg Ala Glu Ile Asp Met Leu Asp Ile Arg Ala His
275 280 285
Phe Lys Arg Leu Tyr Gly Lys Ser Leu Tyr Ser Phe Ile Lys Gly Asp
290 295 300
Thr Ser Gly Asp Tyr Arg Lys Val Leu Leu Val Leu Cys Gly Gly Asp
305 310 315 320
Asp

<210> 11

<211> 1316
<212> DNA
<213> Homo sapiens

<400> 11
agctagacgc cccgaggtcg gagtgaagcg ccgggaccga gccccgtctc ccagggagtc 60
cggggcgcac ggcaccgagg agagcgcggg agccaacctg ggcgcatcat gcgcagggcc 120
cgggacgctg ggccggtcta caccgccgcc tgggtcacgt ggccccggacg ggccggcggc 180
tgccccggcc gggggggcgg ggtcgcgccg gggttgcgct ggacgacgga gagcggcggg 240
cccgcagcgg cctggagcct cccaacccgc gccgcgctgg ccctcgagcg taggacccgc 300
ccctgcccc cccgcgccgg ccccgcgccc ggccgccccgc cccctatata gcgcgcccca 360
gcagggcccg cgccaggccg ccagcctcgg agtgggcgcg ggacagtgcg cggcgccccg 420
cagccaggcc cccgccccgg ccgcatccac ctcctccgcc gcctgcgacc caacgggcgc 480
ccccgccggc cagctcgcgc cgggccccg cggccaccat gaagaaggag gtgtgctccg 540
tggccttcct caaggccgtg ttcgcagagt tcttggccac cctcatcttc gtcttctttg 600
gcctgggctc ggccctcaag tggccgtcgg cgctgcctac catcctgcag atcgcgctgg 660
cgtttggcct ggccataggc acgctggccc aggccctggg acccgtgagc ggcggccaca 720
tcaaccccgc catcaccctg gccctcttgg tgggcaacca gatctcgctg ctccgggctt 780
tcttctacgt ggcggcccag ctggtgggcg ccattgccgg ggctggcatc ctctacggtg 840
tggcaccgct caatgcccgg ggcaatctgg ccgtcaacgc gctcaacaac aacacaacgc 900
agggccaggc catggtggtg gagctgattc tgaccttcca gctggcactc tgcatcttcg 960
cctccactga ctccgccgcc accagccctg tgggctcccc agccctgtcc attggcctgt 1020
ctgtcaccct gggccacctt gtcggaatct acttcactgg ctgctccatg aacccagccc 1080
gctcttttgg ccctgcggtg gtcatgaatc ggttcagccc cgctcactgg gttttctggg 1140
tagggcccat cgtgggggcg gtcctggctg ccatccttta cttctacctg ctcttcccca 1200
actcctgag cctgagtgag cgtgtggcca tcatcaaagg cacgtatgag cctgacgagg 1260
actgggagga gcagcgggaa gagcggaaga agaccatgga gctgaccacc cgctga 1316

<210> 12
<211> 265
<212> PRT
<213> Homo sapiens

<400> 12
Met Lys Lys Glu Val Cys Ser Val Ala Phe Leu Lys Ala Val Phe Ala
 1   5   10   15
Glu Phe Leu Ala Thr Leu Ile Phe Val Phe Phe Gly Leu Gly Ser Ala
            20   25   30
Leu Lys Trp Pro Ser Ala Leu Pro Thr Ile Leu Gln Ile Ala Leu Ala
        35   40   45
Phe Gly Leu Ala Ile Gly Thr Leu Ala Gln Ala Leu Gly Pro Val Ser
    50   55   60
Gly Gly His Ile Asn Pro Ala Ile Thr Leu Ala Leu Leu Val Gly Asn
65   70   75   80
Gln Ile Ser Leu Leu Arg Ala Phe Phe Tyr Val Ala Ala Gln Leu Val
                85   90   95
Gly Ala Ile Ala Gly Ala Gly Ile Leu Tyr Gly Val Ala Pro Leu Asn
            100   105   110
Ala Arg Gly Asn Leu Ala Val Asn Ala Leu Asn Asn Asn Thr Thr Gln
        115   120   125
Gly Gln Ala Met Val Val Glu Leu Ile Leu Thr Phe Gln Leu Ala Leu
    130   135   140
Cys Ile Phe Ala Ser Thr Asp Ser Arg Arg Thr Ser Pro Val Gly Ser
145   150   155   160
Pro Ala Leu Ser Ile Gly Leu Ser Val Thr Leu Gly His Leu Val Gly
                165   170   175
Ile Tyr Phe Thr Gly Cys Ser Met Asn Pro Ala Arg Ser Phe Gly Pro
            180   185   190

```
Ala Val Val Met Asn Arg Phe Ser Pro Ala His Trp Val Phe Trp Val
       195                 200                 205
Gly Pro Ile Val Gly Ala Val Leu Ala Ala Ile Leu Tyr Phe Tyr Leu
   210                 215                 220
Leu Phe Pro Asn Ser Leu Ser Leu Ser Glu Arg Val Ala Ile Ile Lys
225                 230                 235                 240
Gly Thr Tyr Glu Pro Asp Glu Asp Trp Glu Glu Gln Arg Glu Glu Arg
                245                 250                 255
Lys Lys Thr Met Glu Leu Thr Thr Arg
            260                 265
```

<210> 13
<211> 1653
<212> DNA
<213> Homo sapiens

<400> 13

```
acgtccgggg aggggccagg tgagcggcag acccggcacg caggtggggg ccggccgggt 60
ccgtggccag agctgcagag agacaaggcg gcggcggctg ctgtgctggg tgcagtgagg 120
aagaggccct cggtggtgcc catggctggc caggatcctg cgctgagcac gagtcacccg 180
ttctacgacg tggccagaca tggcattctg caggtggcag gggatgaccg ctttggaaga 240
cgtgttgtca cgttcagctg ctgccggatg ccgccctccc acgagctgga ccaccagcgg 300
ctgctggagt atttgaagta cacactggac caatacgttg agaacgatta taccatcgtc 360
tatttccact acgggctgaa cagccggaac aagcctccca tgggctggct ccagagcgca 420
tacaaggagt tcgatagggaa agacgggggat ctcactatgt ggcccaggct ggtctcgaac 480
tccaagctca agcgatcctc ccacctcagc ctcccaaagt actgggatta caggtacaag 540
aagaacttga aggcctcta cgtggtgcac cccaccagct tcatcaaggt cctgtggaac 600
atcttgaagc cctcatcag tcacaagttt gggaagaaag tcatctattt caactacctg 660
agtgagctcc acgaacacct taaatacgac cagctggtca tccctcccga agttttgcgg 720
tacgatgaga agctccagag cctgcacgag ggccggacgc cgcctcctac caagacacca 780
ccgccgcggc ccccgctgcc cacacagcag tttggcgtca gtctgcaata cctcaaagac 840
aaaaatcaag gcgaactcat ccccctgtg ctgaggttca cagtgacgta cctgagagag 900
aaaggcctgc gcaccgaggg cctgttccgg agatccgcca gcgtgcagac cgtccgcgag 960
atccagaggc tctacaacca agggaagccc gtgaactttg acgactacgg ggacattcac 1020
atccctgccg tgatcctgaa gaccttcctg cgagagctgc cccagccgct tctgaccttc 1080
caggcctacg agcagattct cgggatcacc tgtgtggaga gcagcctgcg tgtcactggc 1140
tgccgccaga tcttacggag cctcccagag cacaactacg tcgtcctccg ctacctcatg 1200
ggctttctgc atgcggtgtc ccgggagagc atcttcaaca aaatgaacag ctctaacctg 1260
gcctgtgtct tcgggctgaa tttgatatgg ccatccacgg gggtctcctc cctgagtgcc 1320
cttgtgcccc tgaacatgtt cactgaactg ctgatcgagt actatgaaaa gatcttcagc 1380
accccggagg cacctgggga gcacggcctg gcaccatggg aacagggggag cagggcagcc 1440
cctttgcagg aggctgtgcc acggacacaa gccacgggcc tcaccaagcc taccctacct 1500
ccgagtcccc tgatggcagc cagaagacgt ctctagtgtt gcgaacactc tgtatgtttc 1560
gagctacctc ccacacctgt ctgtgcactt gtatgttttg taaacttggc atctgtaaaa 1620
ataaccagcc attagatgaa ttcagaacct tct                              1653
```

<210> 14
<211> 464
<212> PRT
<213> Homo sapiens

<400> 14

```
Met Ala Gly Gln Asp Pro Ala Leu Ser Thr Ser His Pro Phe Tyr Asp
 1               5                  10                  15
Val Ala Arg His Gly Ile Leu Gln Val Ala Gly Asp Asp Arg Phe Gly
            20                  25                  30
```

<212> DNA
<213> Homo sapiens

<400> 15
atgaggactc tccgcaggtt gaagttcatg agttcgccca gcctcagtga cctgggcaag 60
agagagccgg ccgccgccgc ggacgagcgg ggcacgcagc agccgccggc ctgcgccaac 120
gccacctgga acagcatcca caacggggtg atcgccgtct tccagcgcaa ggggctgccc 180
gaccaggagc tcttcagcct caacgagggc gtccggcagc tgttgaagac agagctgggg 240
tcctttctca cggagtacct gcagaaccag ctgctgacaa aaggcatggt gatcctttcgg 300
gacaagattc gtttctatga gggacagaag ctgctggact cactggcaga gacctgggac 360
ttctttctca gtgacgtgct gcccatgctg caggccatct tctaccccgt gcagggcaag 420
gagccatcgg tgcgccagct ggccctgctg cactttccgga atgccatcac cctcagtgtg 480
aagctagagg atgcgctggc ccgggcccat gcccgtgtgc cccctgccat cgtgcagatg 540
ctgctggtgc tgcaggggct acatgagtcc aggggcgtga ctgaggacta cctgcgcctg 600
gagacgctgg tccagaaggt ggtgtcgcca tacctgggca cctacggcct ccactccagc 660
gaggggccct tcacccattc ctgcatcctg gagctgcaga gagacaaggc ggcggcggct 720
gctgtgctgg gtgcagtgag gaagaggccc tcggtggtgc ccatggctgg ccaggtatcct 780
gcgctgagca cgagtcaccc gttctacgac gtggccagac atggcattct gcaggtggca 840
ggggatgacc gctttggaag acgtgttgtc acgttcagct gctgccggat gccaccctcc 900
cacgagctgg accaccagcg gctgctggag tacaagaaga acttgaaggc cctctacgtg 960
gtgcacccca ccagcttcat caaggtcctg tggaacatct tgaagccctt catcagtcac 1020
aagtttggga agaaagtcat ctatttcaac tacctgagtg agctccacga acaccttaaa 1080
tacgaccagc tggtcatccc tcccgaagtt ttgcggtacg atgagaagct ccagagcctg 1140
cacgagggcc ggacgccgcc tcccaccaag acacccaccgc cgcggccccc gctgcccaca 1200
cagcagtttg gcgtcagtct gcaatacctc aaagacaaaa atcaaggcga actcatcccc 1260
cctgtgctga ggttcacagt gacgtacctg agagagaaag gcctgcgcac cgaggcctg 1320
ttccggagat ccgccagcgt gcagaccgtc cgcgagatcc agaggctcta caaccaaggg 1380
aagcccgtga actttgacga ctacggggac attcacatcc ctgccgtgat cctgaagacc 1440
ttcctgcgag agctgcccca gccgcttctg accttccagg cctacgagca gattctcggg 1500
atcacctgtg tggagacag cctgcgtgtc actggctgcc gccagatctt acggagcctc 1560
ccagagcaca actacgtcgt cctccgctac ctcatgggct tcctgcatgc ggtgtcccgg 1620
gagagcatct tcaacaaaat gaacagctct aacctggcct gtgtcttcgg gctgaatttg 1680
atctggccat cccaggggct ctcctccctg agtgccttg tgcccctgaa catgttcact 1740
gaactgctga tcgagtacta tgaaaagatc ttcagcaccc cggaggcacc tggggagcac 1800
ggcctggcac catgggaaca ggggagcagg gcagccccct tgcaggaggc tgtgccacgg 1860
acacaagcca cgggcctcac caagcctacc ctacctccga gtccctgat ggcagccaga 1920
agacgtctct agtgttgcga acactctgta tatttcgagc tacctcccac acctgtctgt 1980
gcacttgtat gttttgtaaa cttggcatct gtaaaaataa ccagccatta gatgaattca 2040
gaa 2043

<210> 16
<211> 643
<212> PRT
<213> Homo sapiens

<400> 16
Met Arg Thr Leu Arg Arg Leu Lys Phe Met Ser Ser Pro Ser Leu Ser
 1   5   10   15
Asp Leu Gly Lys Arg Glu Pro Ala Ala Ala Ala Asp Glu Arg Gly Thr
            20   25   30
Gln Gln Arg Arg Ala Cys Ala Asn Ala Thr Trp Asn Ser Ile His Asn
        35   40   45
Gly Val Ile Ala Val Phe Gln Arg Lys Gly Leu Pro Asp Gln Glu Leu
    50   55   60
Phe Ser Leu Asn Glu Gly Val Arg Gln Leu Leu Lys Thr Glu Leu Gly
65   70   75   80
Ser Phe Phe Thr Glu Tyr Leu Gln Asn Gln Leu Leu Thr Lys Gly Met
                85   90   95

```
Val Ile Leu Arg Asp Lys Ile Arg Phe Tyr Glu Gly Gln Lys Leu Leu
            100                 105                 110
Asp Ser Leu Ala Glu Thr Trp Asp Phe Phe Phe Ser Asp Val Leu Pro
        115                 120                 125
Met Leu Gln Ala Ile Phe Tyr Pro Val Gln Gly Lys Glu Pro Ser Val
    130                 135                 140
Arg Gln Leu Ala Leu Leu His Phe Arg Asn Ala Ile Thr Leu Ser Val
145                 150                 155                 160
Lys Leu Glu Asp Ala Leu Ala Arg Ala His Ala Arg Val Pro Pro Ala
                165                 170                 175
Ile Val Gln Met Leu Leu Val Leu Gln Gly Val His Glu Ser Arg Gly
            180                 185                 190
Val Thr Glu Asp Tyr Leu Arg Leu Glu Thr Leu Val Gln Lys Val Val
        195                 200                 205
Ser Pro Tyr Leu Gly Thr Tyr Gly Leu His Ser Ser Glu Gly Pro Phe
    210                 215                 220
Thr His Ser Cys Ile Leu Glu Leu Gln Arg Asp Lys Ala Ala Ala Ala
225                 230                 235                 240
Ala Val Leu Gly Ala Val Arg Lys Arg Pro Ser Val Val Pro Met Ala
                245                 250                 255
Gly Gln Asp Pro Ala Leu Ser Thr Ser His Pro Phe Tyr Asp Val Ala
            260                 265                 270
Arg His Gly Ile Leu Gln Val Ala Gly Asp Asp Arg Phe Gly Arg Arg
        275                 280                 285
Val Val Thr Phe Ser Cys Cys Arg Met Pro Pro Ser His Glu Leu Asp
    290                 295                 300
His Gln Arg Leu Leu Glu Tyr Lys Lys Asn Leu Lys Ala Leu Tyr Val
305                 310                 315                 320
Val His Pro Thr Ser Phe Ile Lys Val Leu Trp Asn Ile Leu Lys Pro
                325                 330                 335
Leu Ile Ser His Lys Phe Gly Lys Lys Val Ile Tyr Phe Asn Tyr Leu
            340                 345                 350
Ser Glu Leu His Glu His Leu Lys Tyr Asp Gln Leu Val Ile Pro Pro
        355                 360                 365
Glu Val Leu Arg Tyr Asp Glu Lys Leu Gln Ser Leu His Glu Gly Arg
    370                 375                 380
Thr Pro Pro Pro Thr Lys Thr Pro Pro Pro Arg Pro Pro Leu Pro Thr
385                 390                 395                 400
Gln Gln Phe Gly Val Ser Leu Gln Tyr Leu Lys Asp Lys Asn Gln Gly
                405                 410                 415
Glu Leu Ile Pro Pro Val Leu Arg Phe Thr Val Thr Tyr Leu Arg Glu
            420                 425                 430
Lys Gly Leu Arg Thr Glu Gly Leu Phe Arg Arg Ser Ala Ser Val Gln
        435                 440                 445
Thr Val Arg Glu Ile Gln Arg Leu Tyr Asn Gln Gly Lys Pro Val Asn
    450                 455                 460
Phe Asp Asp Tyr Gly Asp Ile His Ile Pro Ala Val Ile Leu Lys Thr
465                 470                 475                 480
Phe Leu Arg Glu Leu Pro Gln Pro Leu Leu Thr Phe Gln Ala Tyr Glu
                485                 490                 495
Gln Ile Leu Gly Ile Thr Cys Val Glu Ser Ser Leu Arg Val Thr Gly
            500                 505                 510
Cys Arg Gln Ile Leu Arg Ser Leu Pro Glu His Asn Tyr Val Val Leu
        515                 520                 525
Arg Tyr Leu Met Gly Phe Leu His Ala Val Ser Arg Glu Ser Ile Phe
    530                 535                 540
Asn Lys Met Asn Ser Ser Asn Leu Ala Cys Val Phe Gly Leu Asn Leu
545                 550                 555                 560
```

```
Ile Trp Pro Ser Gln Gly Val Ser Ser Leu Ser Ala Leu Val Pro Leu
                565                 570                 575
Asn Met Phe Thr Glu Leu Leu Ile Glu Tyr Tyr Glu Lys Ile Phe Ser
            580                 585                 590
Thr Pro Glu Ala Pro Gly Glu His Gly Leu Ala Pro Trp Glu Gln Gly
        595                 600                 605
Ser Arg Ala Ala Pro Leu Gln Glu Ala Val Pro Arg Thr Gln Ala Thr
    610                 615                 620
Gly Leu Thr Lys Pro Thr Leu Pro Pro Ser Pro Leu Met Ala Ala Arg
625                 630                 635                 640
Arg Arg Leu
```

<210> 17
<211> 2274
<212> DNA
<213> Homo sapiens

<400> 17

```
atgaggactc tccgcaggtt gaagttcatg agttcgccca gcctcagtga cctgggcaag 60
agagagccgg ccgccgccgc ggacgagcgg ggcacgcagc agcgccgggc ctgcgccaac 120
gccacctgga acagcatcca caacggggtg atcgccgtct tccagcgcaa ggggctgccc 180
gaccaggagc tcttcagcct caacgagggc gtccggcagc tgttgaagac agagctgggg 240
tcattcttca cggagtacct gcagaaccag ctgctgacaa aaggcatggt gatccttggg 300
gacaagattc gcttctatga gggacagaag ctgctggact cactggcaga gacctgggac 360
ttattcttca gtgacgtgct gcccatgctg caggccatct tctaccoggt gcagggcaag 420
gagccatcgg tgcgccagct ggccctgctg cacttccgga atgccatcac cctcagtgtg 480
aagctagagg atgcgctggc ccgggcccat gcccgtgtgc cccctgccat cgtgcagatg 540
ctgctggtgc tgcaggaggt acatgagtcc agggggctga ctgaggacta cctgcgcctg 600
gagacgctgg tccagaaggt ggtgtcgcca tacctgggca cctacggcct ccactccagc 660
gaggggccct tcacccattc ctgcatcctg gagctgcaga gagacaaggc ggcggcggct 720
gctgtgctgg gtgcagtgag gaagaggccc tcggtggtgc ccatggctgg ccaggatcct 780
gcgctgagca cgagtcaccc gttctacgac gtggccagac atggcattct gcaggtggca 840
ggggatgacc gctttggaag acgtgttgtc acgttcagct gctgccggat gccaccctcc 900
cacgagctgg accaccagcg gctgctggag tatttgaagt acacactgga ccaatacgtt 960
gagaacgatt ataccatcgt ctatttccac tacgggctga acagccggaa caagccttcc 1020
ctgggctggc tccagagcgc atacaaggag ttcgatagga aagacgggga tctcactatg 1080
tggcccaggc tggtctcgaa ctccaagctc aagcgatcct cccacctcag cctcccaaag 1140
tactgggatt acaggtacaa gaagaacttg aaggccctct acgtggtgca ccccaccagc 1200
ttcatcaagg tcctgtggaa catcttgaag cccctcatca gtcacaagtt tgggaagaaa 1260
gtcatctatt tcaactacct gagtgagctc cacgaacacc ttaaatacga ccagctggtc 1320
atccctcccg aagttttgcg gtacgatgag aagctccaga gcctgcacga gggccggacg 1380
ccgcctccca ccaagacacc accgccgcgg cccccgctgc ccacacagca gtttggctgc 1440
agtctgcaat acctcaaaga caaaaatcaa ggcgaactca tccccctgt gctgaggttc 1500
acagtgacgt acctgagaga gaaaggcctg cgcaccgagg gcctgttccg gagatccgcc 1560
agcgtgcaga ccgtccgcga gatccagagg ctctacaacc aagggaagcc cgtgaacttt 1620
gacgactacg gggacattca catccctgcc gtgatcctga agaccttcct gcgagagctg 1680
ccccagccgc ttctgacctt ccaggcctac gagcagattc tcgggatcac ctgtgtggag 1740
agcagcctgc gtgtcactgg ctgccgccag atcttacgga gcctcccaga gcacaactac 1800
gtcgtcctcc gctacctcat gggcttcctg catgcggtgt cccggagagg catcttcaac 1860
aaaatgaaca gctctaacct ggcctgtgtc ttcgggctga atttgatctg gccatcccag 1920
ggggtctcct ccctgagtgc ccttgtgccc ctgaacatgt tcactgaact gctgatcgag 1980
tactatgaaa agatcttcag caccccggag gcacctgggg agcacggcct ggcaccatgg 2040
gaacagggga gcagggcagc ccctttgcag gaggctgtgc cacggacaca agccacgggc 2100
ctcaccaagc ctacctacc tccgagtccc ctgatggcag ccagaagacg tctctagtgt 2160
tgcgaacact ctgtatattt cgagctacct cccacacctg tctgtgcact tgtatgtttt 2220
gtaaacttgg catctgtaaa aataaccagc cattagatga attcagaacc ttct       2274
```

<210> 18
<211> 751
<212> PRT
<213> Homo sapiens

<220>
<221> VARIANT
<222> (1)...(751)
<223> Xaa = Any Amino Acid

<400> 18
Met Arg Thr Leu Arg Arg Leu Lys Phe Met Ser Ser Pro Ser Leu Ser
1 5 10 15
Asp Leu Gly Lys Arg Glu Pro Ala Ala Ala Ala Asp Glu Arg Gly Thr
20 25 30
Gln Gln Arg Arg Ala Cys Ala Asn Ala Thr Trp Asn Ser Ile His Asn
35 40 45
Gly Val Ile Ala Val Phe Gln Arg Lys Gly Leu Pro Asp Gln Glu Leu
50 55 60
Phe Ser Leu Asn Glu Gly Val Arg Gln Leu Leu Lys Thr Glu Leu Gly
65 70 75 80
Ser Phe Phe Thr Glu Tyr Leu Gln Asn Gln Leu Leu Thr Lys Gly Met
85 90 95
Val Ile Leu Arg Asp Lys Ile Arg Phe Tyr Glu Gly Gln Lys Leu Leu
100 105 110
Asp Ser Leu Ala Glu Thr Trp Asp Phe Phe Phe Ser Asp Val Leu Pro
115 120 125
Met Leu Gln Ala Ile Phe Tyr Pro Val Gln Gly Lys Glu Pro Ser Val
130 135 140
Arg Gln Leu Ala Leu Leu His Phe Arg Asn Ala Ile Thr Leu Ser Val
145 150 155 160
Lys Leu Glu Asp Ala Leu Ala Arg Ala His Ala Arg Val Pro Pro Ala
165 170 175
Ile Val Gln Met Leu Leu Val Leu Gln Gly Val His Glu Ser Arg Gly
180 185 190
Val Thr Glu Asp Tyr Leu Arg Leu Glu Thr Leu Val Gln Lys Val Val
195 200 205
Ser Pro Tyr Leu Gly Thr Tyr Gly Leu His Ser Ser Glu Gly Pro Phe
210 215 220
Thr His Ser Cys Ile Leu Glu Leu Gln Arg Asp Lys Ala Ala Ala Ala
225 230 235 240
Ala Val Leu Gly Ala Val Arg Lys Arg Pro Ser Val Val Pro Met Ala
245 250 255
Gly Gln Asp Pro Ala Leu Ser Thr Ser His Pro Phe Tyr Asp Val Ala
260 265 270
Arg His Gly Ile Leu Gln Val Ala Gly Asp Asp Arg Phe Gly Arg Arg
275 280 285
Val Val Thr Phe Ser Cys Cys Arg Met Pro Pro Ser His Glu Leu Asp
290 295 300
His Gln Arg Leu Leu Glu Tyr Leu Lys Tyr Thr Leu Asp Gln Tyr Val
305 310 315 320
Glu Asn Asp Tyr Thr Ile Val Tyr Phe His Tyr Gly Leu Asn Ser Arg
325 330 335
Asn Lys Pro Ser Leu Gly Trp Leu Gln Ser Ala Tyr Lys Glu Phe Asp
340 345 350
Arg Lys Asp Gly Asp Leu Thr Met Trp Pro Arg Leu Val Ser Asn Ser
355 360 365

```
Lys Leu Lys Arg Ser Ser His Leu Ser Leu Pro Lys Tyr Trp Asp Tyr
    370                 375                 380
Arg Tyr Lys Lys Asn Leu Lys Ala Leu Tyr Val Val His Pro Thr Ser
385                 390                 395                 400
Phe Ile Lys Val Leu Trp Asn Ile Leu Lys Pro Leu Ile Ser His Lys
                405                 410                 415
Phe Gly Lys Lys Val Ile Tyr Phe Asn Tyr Leu Ser Glu Leu His Glu
            420                 425                 430
His Leu Lys Tyr Asp Gln Leu Val Ile Pro Pro Glu Val Leu Arg Tyr
        435                 440                 445
Asp Glu Lys Leu Gln Ser Leu His Glu Gly Arg Thr Pro Pro Pro Thr
    450                 455                 460
Lys Thr Pro Pro Pro Arg Pro Pro Leu Pro Thr Gln Gln Phe Gly Val
465                 470                 475                 480
Ser Leu Gln Tyr Leu Lys Asp Lys Asn Gln Gly Glu Leu Ile Pro Pro
                485                 490                 495
Val Leu Arg Phe Thr Val Thr Tyr Leu Arg Glu Lys Gly Leu Arg Thr
            500                 505                 510
Glu Gly Leu Phe Arg Arg Ser Ala Ser Val Gln Thr Val Arg Glu Ile
        515                 520                 525
Gln Arg Leu Tyr Asn Gln Gly Lys Pro Val Asn Phe Asp Asp Tyr Gly
    530                 535                 540
Asp Ile His Ile Pro Ala Val Ile Leu Lys Thr Phe Leu Arg Glu Leu
545                 550                 555                 560
Pro Gln Pro Leu Leu Thr Phe Gln Ala Tyr Glu Gln Ile Leu Gly Ile
                565                 570                 575
Thr Cys Val Glu Ser Ser Leu Arg Val Thr Gly Cys Arg Gln Ile Leu
            580                 585                 590
Arg Ser Leu Pro Glu His Asn Tyr Val Val Leu Arg Tyr Leu Met Gly
        595                 600                 605
Phe Leu His Ala Val Ser Arg Glu Ser Ile Phe Asn Lys Met Asn Ser
    610                 615                 620
Ser Asn Leu Ala Cys Val Phe Gly Leu Asn Leu Ile Trp Pro Ser Gln
625                 630                 635                 640
Gly Val Ser Ser Leu Ser Ala Leu Val Pro Leu Asn Met Phe Thr Glu
                645                 650                 655
Leu Leu Ile Glu Tyr Tyr Glu Lys Ile Phe Ser Thr Pro Glu Ala Pro
            660                 665                 670
Gly Glu His Gly Leu Ala Pro Trp Glu Gln Gly Ser Arg Ala Ala Pro
        675                 680                 685
Leu Gln Glu Ala Val Pro Arg Thr Gln Ala Thr Gly Leu Thr Lys Pro
    690                 695                 700
Thr Leu Pro Pro Ser Pro Leu Met Ala Ala Arg Arg Arg Leu Xaa Cys
705                 710                 715                 720
Cys Glu His Ser Val Tyr Phe Glu Leu Pro Pro Thr Pro Val Cys Ala
                725                 730                 735
Leu Val Cys Phe Val Asn Leu Ala Ser Val Lys Ile Thr Ser His
            740                 745                 750
```

<210> 19
<211> 718
<212> PRT
<213> Homo sapiens

<400> 19

```
Met Arg Thr Leu Arg Arg Leu Lys Pha Met Ser Ser Pro Ser Leu Ser
 1               5                  10                  15
```

```
Asp Leu Gly Lys Arg Glu Pro Ala Ala Ala Ala Asp Glu Arg Gly Thr
            20                  25                  30
Gln Gln Arg Arg Ala Cys Ala Asn Ala Thr Trp Asn Ser Ile His Asn
        35                  40                  45
Gly Val Ile Ala Val Phe Gln Arg Lys Gly Leu Pro Asp Gln Glu Leu
    50                  55                  60
Phe Ser Leu Asn Glu Gly Val Arg Gln Leu Leu Lys Thr Glu Leu Gly
65                  70                  75                  80
Ser Phe Phe Thr Glu Tyr Leu Gln Asn Gln Leu Leu Thr Lys Gly Met
                85                  90                  95
Val Ile Leu Arg Asp Lys Ile Arg Phe Tyr Glu Gly Gln Lys Leu Leu
            100                 105                 110
Asp Ser Leu Ala Glu Thr Trp Asp Phe Phe Phe Ser Asp Val Leu Pro
        115                 120                 125
Met Leu Gln Ala Ile Phe Tyr Pro Val Gln Gly Lys Glu Pro Ser Val
    130                 135                 140
Arg Gln Leu Ala Leu Leu His Phe Arg Asn Ala Ile Thr Leu Ser Val
145                 150                 155                 160
Lys Leu Glu Asp Ala Leu Ala Arg Ala His Ala Arg Val Pro Pro Ala
                165                 170                 175
Ile Val Gln Met Leu Leu Val Leu Gln Gly Val His Glu Ser Arg Gly
            180                 185                 190
Val Thr Glu Asp Tyr Leu Arg Leu Glu Thr Leu Val Gln Lys Val Val
        195                 200                 205
Ser Pro Tyr Leu Gly Thr Tyr Gly Leu His Ser Ser Glu Gly Pro Phe
    210                 215                 220
Thr His Ser Cys Ile Leu Glu Leu Gln Arg Asp Lys Ala Ala Ala Ala
225                 230                 235                 240
Ala Val Leu Gly Ala Val Arg Lys Arg Pro Ser Val Val Pro Met Ala
                245                 250                 255
Gly Gln Asp Pro Ala Leu Ser Thr Ser His Pro Phe Tyr Asp Val Ala
            260                 265                 270
Arg His Gly Ile Leu Gln Val Ala Gly Asp Asp Arg Phe Gly Arg Arg
        275                 280                 285
Val Val Thr Phe Ser Cys Cys Arg Met Pro Pro Ser His Glu Leu Asp
    290                 295                 300
His Gln Arg Leu Leu Glu Tyr Leu Lys Tyr Thr Leu Asp Gln Tyr Val
305                 310                 315                 320
Glu Asn Asp Tyr Thr Ile Val Tyr Phe His Tyr Gly Leu Asn Ser Arg
                325                 330                 335
Asn Lys Pro Ser Leu Gly Trp Leu Gln Ser Ala Tyr Lys Glu Phe Asp
            340                 345                 350
Arg Lys Asp Gly Asp Leu Thr Met Trp Pro Arg Leu Val Ser Asn Ser
        355                 360                 365
Lys Leu Lys Arg Ser Ser His Leu Ser Leu Pro Lys Tyr Trp Asp Tyr
    370                 375                 380
Arg Tyr Lys Lys Asn Leu Lys Ala Leu Tyr Val Val His Pro Thr Ser
385                 390                 395                 400
Phe Ile Lys Val Leu Trp Asn Ile Leu Lys Pro Leu Ile Ser His Lys
                405                 410                 415
Phe Gly Lys Lys Val Ile Tyr Phe Asn Tyr Leu Ser Glu Leu His Glu
            420                 425                 430
His Leu Lys Tyr Asp Gln Leu Val Ile Pro Pro Glu Val Leu Arg Tyr
        435                 440                 445
Asp Glu Lys Leu Gln Ser Leu His Glu Gly Arg Thr Pro Pro Pro Thr
    450                 455                 460
Lys Thr Pro Pro Pro Arg Pro Pro Leu Pro Thr Gln Gln Phe Gly Val
465                 470                 475                 480
```

```
Ser Leu Gln Tyr Leu Lys Asp Lys Asn Gln Gly Glu Leu Ile Pro Pro
                485                 490                 495
Val Leu Arg Phe Thr Val Thr Tyr Leu Arg Glu Lys Gly Leu Arg Thr
            500                 505                 510
Glu Gly Leu Phe Arg Arg Ser Ala Ser Val Gln Thr Val Arg Glu Ile
        515                 520                 525
Gln Arg Leu Tyr Asn Gln Gly Lys Pro Val Asn Phe Asp Asp Tyr Gly
    530                 535                 540
Asp Ile His Ile Pro Ala Val Ile Leu Lys Thr Phe Leu Arg Glu Leu
545                 550                 555                 560
Pro Gln Pro Leu Leu Thr Phe Gln Ala Tyr Glu Gln Ile Leu Gly Ile
                565                 570                 575
Thr Cys Val Glu Ser Ser Leu Arg Val Thr Gly Cys Arg Gln Ile Leu
            580                 585                 590
Arg Ser Leu Pro Glu His Asn Tyr Val Val Leu Arg Tyr Leu Met Gly
        595                 600                 605
Phe Leu His Ala Val Ser Arg Glu Ser Ile Phe Asn Lys Met Asn Ser
    610                 615                 620
Ser Asn Leu Ala Cys Val Phe Gly Leu Asn Leu Ile Trp Pro Ser Gln
625                 630                 635                 640
Gly Val Ser Ser Leu Ser Ala Leu Val Pro Leu Asn Met Phe Thr Glu
                645                 650                 655
Leu Leu Ile Glu Tyr Tyr Glu Lys Ile Phe Ser Thr Pro Glu Ala Pro
            660                 665                 670
Gly Glu His Gly Leu Ala Pro Trp Glu Gln Gly Ser Arg Ala Ala Pro
        675                 680                 685
Leu Gln Glu Ala Val Pro Arg Thr Gln Ala Thr Gly Leu Thr Lys Pro
    690                 695                 700
Thr Leu Pro Pro Ser Pro Leu Met Ala Ala Arg Arg Arg Leu
705                 710                 715
```

<210> 20
<211> 1431
<212> DNA
<213> Homo sapiens

<400> 20

```
acgtccgggg aggggccagg tgagcggcag acccggcacg caggtggggg ccggcggggt 60
ccgtggccag agctgcagag agacaaggcg gcggcggctg ctgtgctggg tgcagtgagg 120
aagaggccct cggtggtgcc catggctggc caggatcctg cgctgagcac gagtcacccg 180
ttctacgacg tggccagaca tggcattctg caggtggcag gggatgaccg ctttggaaga 240
cgtgttgtca cgttcagctg ctgccggatg ccaccctccc acgagctgga ccaccagcgg 300
ctgctggaca ggtacaagaa gaacttgaag gccctctacg tggtgcaccc caccagcttc 360
atcaaggtcc tgtggaacat cttgaagccc ctcatcagtc acaagtttgg gaagaaagtc 420
atctatttca actacctgag tgagctccac gaacacctta aatacgacca gctggtcatc 480
cctcccgaag ttttgcggta cgatgagaag ctccagagcc tgcacgaggg ccggacgccg 540
cctcccacca agacaccacc gccgcggccc ccgctgccca cacagcagtt tggcgtcagt 600
ctgcaatacc tcaaagacaa aaatcaaggc gaactcatcc cccctgtgct gaggttcaca 660
gtgacgtacc tgagagagaa aggcctgcgc accgagggcc tgttccggag atccgccagc 720
gtgcagaccg tccgcgagat ccagaggctc tacaaccaag ggaagcccgt gaactttgac 780
gactacgggg acattcacat ccctgccgtg atcctgaaga ccttcctgcg agagctgccc 840
cagccgcttc tgaccttcca ggcctacgag cagattctcg ggatcacctg tgtggagagc 900
agcctgcgtg tcactggctg ccgccagatc ttacggagcc tcccagagca caactacgtc 960
gtcctccgct acctcatggg cttcctgcat gcggtgtccc gggagagcat cttcaacaaa 1020
atgaacagct ctaacctggc ctgtgtcttc gggctgaatt tgatctggcc atcccagggg 1080
gtctcctccc tgagtgccct tgtgcccctg aacatgttca ctgaactgct gatcgagtac 1140
tatgaaaaga tcttcagcac cccggaggca cctggggagc acggcctggc accatgggaa 1200
```

cagggagca gggcagcccc tttgcaggag gctgtgccac ggacacaagc cacgggcctc 1260
accaagccta ccctacctcc gagtcccctg atggcagcca gaagacgtct ctagtgttgc 1320
gaacactctg tatatttcga gctacctccc acacctgtct gtgcacttgt atgttttgta 1380
aacttggcat ctgtaaaaat aaccagccat tagatgaatt cagaaccttc t 1431

<210> 21
<211> 390
<212> PRT
<213> Homo sapiens

<400> 21
Met Ala Gly Gln Asp Pro Ala Leu Ser Thr Ser His Pro Phe Tyr Asp
1 5 10 15
Val Ala Arg His Gly Ile Leu Gln Val Ala Gly Asp Asp Arg Phe Gly
20 25 30
Arg Arg Val Val Thr Phe Ser Cys Cys Arg Met Pro Pro Ser His Glu
35 40 45
Leu Asp His Gln Arg Leu Leu Asp Arg Tyr Lys Lys Asn Leu Lys Ala
50 55 60
Leu Tyr Val Val His Pro Thr Ser Phe Ile Lys Val Leu Trp Asn Ile
65 70 75 80
Leu Lys Pro Leu Ile Ser His Lys Phe Gly Lys Lys Val Ile Tyr Phe
85 90 95
Asn Tyr Leu Ser Glu Leu His Glu His Leu Lys Tyr Asp Gln Leu Val
100 105 110
Ile Pro Pro Glu Val Leu Arg Tyr Asp Glu Lys Leu Gln Ser Leu His
115 120 125
Glu Gly Arg Thr Pro Pro Pro Thr Lys Thr Pro Pro Pro Arg Pro Pro
130 135 140
Leu Pro Thr Gln Gln Phe Gly Val Ser Leu Gln Tyr Leu Lys Asp Lys
145 150 155 160
Asn Gln Gly Glu Leu Ile Pro Pro Val Leu Arg Phe Thr Val Thr Tyr
165 170 175
Leu Arg Glu Lys Gly Leu Arg Thr Glu Gly Leu Phe Arg Arg Ser Ala
180 185 190
Ser Val Gln Thr Val Arg Glu Ile Gln Arg Leu Tyr Asn Gln Gly Lys
195 200 205
Pro Val Asn Phe Asp Asp Tyr Gly Asp Ile His Ile Pro Ala Val Ile
210 215 220
Leu Lys Thr Phe Leu Arg Glu Leu Pro Gln Pro Leu Leu Thr Phe Gln
225 230 235 240
Ala Tyr Glu Gln Ile Leu Gly Ile Thr Cys Val Glu Ser Ser Leu Arg
245 250 255
Val Thr Gly Cys Arg Gln Ile Leu Arg Ser Leu Pro Glu His Asn Tyr
260 265 270
Val Val Leu Arg Tyr Leu Met Gly Phe Leu His Ala Val Ser Arg Glu
275 280 285
Ser Ile Phe Asn Lys Met Asn Ser Ser Asn Leu Ala Cys Val Phe Gly
290 295 300
Leu Asn Leu Ile Trp Pro Ser Gln Gly Val Ser Ser Leu Ser Ala Leu
305 310 315 320
Val Pro Leu Asn Met Phe Thr Glu Leu Leu Ile Glu Tyr Tyr Glu Lys
325 330 335
Ile Phe Ser Thr Pro Glu Ala Pro Gly Glu His Gly Leu Ala Pro Trp
340 345 350
Glu Gln Gly Ser Arg Ala Ala Pro Leu Gln Glu Ala Val Pro Arg Thr
355 360 365
Gln Ala Thr Gly Leu Thr Lys Pro Thr Leu Pro Pro Ser Pro Leu Met

```
    370                 375                 380
Ala Ala Arg Arg Arg Leu
385                 390


<210> 22
<211> 2019
<212> DNA
<213> Homo sapiens

<400> 22
atgaggactc tccgcaggtt gaagttcatg agttcgccca gcctcagtga cctgggcaag 60
agagagccgg ccgccgccgc ggacgagcgg ggcacgcagc agcgccgggc ctgcgccaac 120
gccacctgga acagcatcca caacgggctg atcgccgtct tccagcgcaa ggggctgccc 180
gaccaggagc tcttcagcct caacgagggc gtccggcagc tgttgaagac agagctgggg 240
tccttcttca cggagtacct gcagaaccag ctgctgacaa aaggcatggt gatccttcgg 300
gacaagattc gcttctatga gggacagaag ctgctggact cactggcaga gactgggac 360
ttcttcttca gtgacgtgct gcccatgctg caggccatct tctaccacgt gcagggcaag 420
gagccatcgg tgccgcagct ggccctgctg caacttccgga atgccatcac cctcagtgtg 480
aagctagagg atgcgctggc ccgggcccat gcccgtgtgc ccctgccat cgtgcagatg 540
ctgctggtgc tgcaggcggt acatgagtcc agggcgctga ctgagcacta cctgcgcctg 600
gagacgctgg tccagaaggt ggtgtcgcca tacctgggca cctacggcct ccactccagc 660
gaggggccct tcacccattc ctgcatcctg gagctgcaga gagacaaggc ggcggcggct 720
gctgtgctgg gtgcagtgag gaagaggccc tcggtggtgc ccatggctgg ccaggatcct 780
gcgctgagca cgagtcaccc gttctacgac gtggccagac atggcattct gcaggtggca 840
ggggatgacc gctttggaag acgtgttgtc acgttcagct gctgcggat gccacctcc 900
cacgagctgg accaccagcg gctgctggag tatttgaagt acacactgga ccaatacgtt 960
gagaacgatt ataccatcgt ctatttccac tacgggctga acagcggga caagccttcc 1020
ctgggctggc tccagagcgc atacaaggag ttcgatagga aagacgggga tctcactatg 1080
tggcccaggc tggtctgaa ctccaagctc aagcgatcct ccatcctcag cctcccaaag 1140
tactgggatt acaggtacaa gaagaacttg aaggccctct acgtggtgca cccaccagc 1200
ttcatcaagg tcctgtggaa catcttgaag ccccatcata gtcacagttt tgggaagaaa 1260
gtcatctatt tcaactacct gagtgagctc cacgaacacc ttaaatacga ccagctggtc 1320
atccctccag aagttttgcg gtacgatgag aagctccaga gcctgcacga gggccggacg 1380
ccgcctccca ccaagacacc accgccgcgg cccccgctgc ccacacagca gtttggcgtc 1440
agtctgcaat acctcaaaga caaaaatcaa ggcgaactca tccccctgt gctgaggttc 1500
acagtgacgt acctgagaga gaaaggcctc ccagagcaca actacgtcgt cctccgctac 1560
ctcatgggct tcgtcgatgc ggtgtccogg gagagcatct tcaacgaaat gaacagctct 1620
aacctggcct gtgtcttcgg gcigaatttg atctggccat cccaggggt ctcctccctg 1680
agtgccttg tgcccctgaa catgttcact gaactgctga tcgagtacta tgaaaagatc 1740
ttcagcaccc cggaggcacc tggggagcac ggcctggcac catgggaaca ggggagcagg 1800
gcagcccctt tgcaggaggc tgtgccacgg acacaagcca cgggcctcac caagcctacc 1860
ctacctccga gtccccigat ggcagccaga agacgtctct agtgttgcga acactctgta 1920
tatttcgagc tacctcccac acctgtctgt gcacttgtat gttttgtaaa cttggcatct 1980
gtaaaaataa ccagccatta gatgaattca gaaccttct                       2019

<210> 23
<211> 633
<212> PRT
<213> Homo sapiens

<400> 23
Met Arg Thr Leu Arg Arg Leu Lys Phe Met Ser Ser Pro Ser Leu Ser
 1               5                  10                  15
Asp Leu Gly Lys Arg Glu Pro Ala Ala Ala Ala Asp Glu Arg Gly Thr
            20                  25                  30
Gln Gln Arg Arg Ala Cys Ala Asn Ala Thr Trp Asn Ser Ile His Asn
        35                  40                  45
```

```
Gly Val Ile Ala Val Phe Gln Arg Lys Gly Leu Pro Asp Gln Glu Leu
    50                  55                  60
Phe Ser Leu Asn Glu Gly Val Arg Gln Leu Leu Lys Thr Glu Leu Gly
65                  70                  75                  80
Ser Phe Phe Thr Glu Tyr Leu Gln Asn Gln Leu Leu Thr Lys Gly Met
                85                  90                  95
Val Ile Leu Arg Asp Lys Ile Arg Phe Tyr Glu Gly Gln Lys Leu Leu
            100                 105                 110
Asp Ser Leu Ala Glu Thr Trp Asp Phe Phe Phe Ser Asp Val Leu Pro
        115                 120                 125
Met Leu Gln Ala Ile Phe Tyr Pro Val Gln Gly Lys Glu Pro Ser Val
    130                 135                 140
Arg Gln Leu Ala Leu Leu His Phe Arg Asn Ala Ile Thr Leu Ser Val
145                 150                 155                 160
Lys Leu Glu Asp Ala Leu Ala Arg Ala His Ala Arg Val Pro Pro Ala
                165                 170                 175
Ile Val Gln Met Leu Leu Val Leu Gln Gly Val His Glu Ser Arg Gly
            180                 185                 190
Val Thr Glu Asp Tyr Leu Arg Leu Glu Thr Leu Val Gln Lys Val Val
        195                 200                 205
Ser Pro Tyr Leu Gly Thr Tyr Gly Leu His Ser Ser Glu Gly Pro Phe
    210                 215                 220
Thr His Ser Cys Ile Leu Glu Leu Gln Arg Asp Lys Ala Ala Ala Ala
225                 230                 235                 240
Ala Val Leu Gly Ala Val Arg Lys Arg Pro Ser Val Val Pro Met Ala
                245                 250                 255
Gly Gln Asp Pro Ala Leu Ser Thr Ser His Pro Phe Tyr Asp Val Ala
            260                 265                 270
Arg His Gly Ile Leu Gln Val Ala Gly Asp Asp Arg Phe Gly Arg Arg
        275                 280                 285
Val Val Thr Phe Ser Cys Cys Arg Met Pro Pro Ser His Glu Leu Asp
    290                 295                 300
His Gln Arg Leu Leu Glu Tyr Leu Lys Tyr Thr Leu Asp Gln Tyr Val
305                 310                 315                 320
Glu Asn Asp Tyr Thr Ile Val Tyr Phe His Tyr Gly Leu Asn Ser Arg
                325                 330                 335
Asn Lys Pro Ser Leu Gly Trp Leu Gln Ser Ala Tyr Lys Glu Phe Asp
            340                 345                 350
Arg Lys Asp Gly Asp Leu Thr Met Trp Pro Arg Leu Val Ser Asn Ser
        355                 360                 365
Lys Leu Lys Arg Ser Ser His Leu Ser Leu Pro Lys Tyr Trp Asp Tyr
    370                 375                 380
Arg Tyr Lys Lys Asn Leu Lys Ala Leu Tyr Val Val His Pro Thr Ser
385                 390                 395                 400
Phe Ile Lys Val Leu Trp Asn Ile Leu Lys Pro Leu Ile Ser His Lys
                405                 410                 415
Phe Gly Lys Lys Val Ile Tyr Phe Asn Tyr Leu Ser Glu Leu His Glu
            420                 425                 430
His Leu Lys Tyr Asp Gln Leu Val Ile Pro Pro Glu Val Leu Arg Tyr
        435                 440                 445
Asp Glu Lys Leu Gln Ser Leu His Glu Gly Arg Thr Pro Pro Pro Thr
    450                 455                 460
Lys Thr Pro Pro Pro Arg Pro Pro Leu Pro Thr Gln Gln Phe Gly Val
465                 470                 475                 480
Ser Leu Gln Tyr Leu Lys Asp Lys Asn Gln Gly Glu Leu Ile Pro Pro
                485                 490                 495
Val Leu Arg Phe Thr Val Thr Tyr Leu Arg Glu Lys Gly Leu Pro Glu
            500                 505                 510
```

```
His Asn Tyr Val Val Leu Arg Tyr Leu Met Gly Phe Leu His Ala Val
        515                 520                 525
Ser Arg Glu Ser Ile Phe Asn Lys Met Asn Ser Ser Asn Leu Ala Cys
    530                 535                 540
Val Phe Gly Leu Asn Leu Ile Trp Pro Ser Gln Gly Val Ser Ser Leu
545                 550                 555                 560
Ser Ala Leu Val Pro Leu Asn Met Phe Thr Glu Leu Leu Ile Glu Tyr
                565                 570                 575
Tyr Glu Lys Ile Phe Ser Thr Pro Glu Ala Pro Gly Glu His Gly Leu
            580                 585                 590
Ala Pro Trp Glu Gln Gly Ser Arg Ala Ala Pro Leu Gln Glu Ala Val
        595                 600                 605
Pro Arg Thr Gln Ala Thr Gly Leu Thr Lys Pro Thr Leu Pro Pro Ser
    610                 615                 620
Pro Leu Met Ala Ala Arg Arg Arg Leu
625                 630
```

<210> 24
<211> 1398
<212> DNA
<213> Homo sapiens

<400> 24

```
acgtccgggg aggggccagg tgagcggcag acccggcacg caggtgggcg ccggcggggt 60
ccgtggccag agctgcagag agacaaggcg gcggcggctg ctgtgctggg tgcagtgagg 120
aagaggccct cggtggtgcc catggctggc caggatcctg cgctgagcac gagtcacccg 180
ttctacgacg tggccagaca tggcattctg caggtggcag gggatgaccg ctttggaaga 240
cgtgttgtca cgttcagctg ctgccggatg ccaccctccc acgagctgga ccaccagcgg 300
ctgctggagt atttgaagta cacactggac caatacgttg agaacgatta taccatcgtc 360
tatttccact acgggctgaa cagccggaac aagccttccc tgggctggct ccagagcgca 420
tacaaggagt tcgataggaa agacggggat ctcactatgt ggcccaggct ggtctcgaac 480
tccaagctca agcgatcctc ccacctcagc ctcccaaagt actgggatta caggtacaag 540
aagaacttga aggcctctca cgtggtgcac cccaccagct tcatcaaggt cctgtggaac 600
atcttgaagc ccctcatcag tcacaagttt gggaagaaag tcatctattt caactacctg 660
agtgagctcc acgaacacct taaatacgac cagctggtca tccctcccga agttttgcgg 720
tacgatgaga agctccagag cctgcacgag ggccggacgc cgcctcccac caagacacca 780
ccgccgcggc ccccgctgcc cacacagcag tttggcgtca gtctgcaata cctcaaagac 840
aaaaatcaag gcgaactcat ccccctgtg ctgaggttca cagtgacgta cctgagagag 900
aaaggcctcc caqagcacaa ctacgtcgtc ctccgctacc tcatgggctt cctgcatgcg 960
gtgtcccggg agagcatctt caacaaaatg aacagctcta acctggcctg tgtcttcggg 1020
ctgaatttga tctggccatc ccaggggtc tcctccctga gtgcccttgt gccctgaac 1080
atgttcactg aactgctgat cgagtactat gaaaagatct tcagcacccc ggaggcacct 1140
ggggagcacg gcctggcacc atgggaacag gggagcaggg cagccccttt gcaggaggct 1200
gtgccacgga cacaagccac gggcctcacc aagcctaccc tacctccgag tcccctgatg 1260
gcagccagaa gacgtcttca gtgttgcgaa cactctgtat atttcgagct acctcccaca 1320
cctgtctgtg cacttgtatg ttttgtaaac ttggcatctg taaaaataac cagccattag 1380
atgaattcag aaccttct                                              1398
```

<210> 25
<211> 379
<212> PRT
<213> Homo sapiens

<400> 25

```
Met Ala Gly Gln Asp Pro Ala Leu Ser Thr Ser His Pro Phe Tyr Asp
 1               5                  10                  15
Val Ala Arg His Gly Ile Leu Gln Val Ala Gly Asp Asp Arg Phe Gly
```

```
            20                  25                  30
Arg Arg Val Val Thr Phe Ser Cys Cys Arg Met Pro Pro Ser His Glu
        35                  40                  45
Leu Asp His Gln Arg Leu Leu Glu Tyr Leu Lys Tyr Thr Leu Asp Gln
    50                  55                  60
Tyr Val Glu Asn Asp Tyr Thr Ile Val Tyr Phe His Tyr Gly Leu Asn
65                  70                  75                  80
Ser Arg Asn Lys Pro Ser Leu Gly Trp Leu Gln Ser Ala Tyr Lys Glu
                85                  90                  95
Phe Asp Arg Lys Asp Gly Asp Leu Thr Met Trp Pro Arg Leu Val Ser
            100                 105                 110
Asn Ser Lys Leu Lys Arg Ser Ser His Leu Ser Leu Pro Lys Tyr Trp
        115                 120                 125
Asp Tyr Arg Tyr Lys Lys Asn Leu Lys Ala Leu Tyr Val Val His Pro
    130                 135                 140
Thr Ser Phe Ile Lys Val Leu Trp Asn Ile Leu Lys Pro Leu Ile Ser
145                 150                 155                 160
His Lys Phe Gly Lys Lys Val Ile Tyr Phe Asn Tyr Leu Ser Glu Leu
                165                 170                 175
His Glu His Leu Lys Tyr Asp Gln Leu Val Ile Pro Pro Glu Val Leu
            180                 185                 190
Arg Tyr Asp Glu Lys Leu Gln Ser Leu His Glu Gly Arg Thr Pro Pro
        195                 200                 205
Pro Thr Lys Thr Pro Pro Pro Arg Pro Pro Leu Pro Thr Gln Gln Phe
    210                 215                 220
Gly Val Ser Leu Gln Tyr Leu Lys Asp Lys Asn Gln Gly Glu Leu Ile
225                 230                 235                 240
Pro Pro Val Leu Arg Phe Thr Val Thr Tyr Leu Arg Glu Lys Gly Leu
                245                 250                 255
Pro Glu His Asn Tyr Val Val Leu Arg Tyr Leu Met Gly Phe Leu His
            260                 265                 270
Ala Val Ser Arg Glu Ser Ile Phe Asn Lys Met Asn Ser Ser Asn Leu
        275                 280                 285
Ala Cys Val Phe Gly Leu Asn Leu Ile Trp Pro Ser Gln Gly Val Ser
    290                 295                 300
Ser Leu Ser Ala Leu Val Pro Leu Asn Met Phe Thr Glu Leu Leu Ile
305                 310                 315                 320
Glu Tyr Tyr Glu Lys Ile Phe Ser Thr Pro Glu Ala Pro Gly Glu His
                325                 330                 335
Gly Leu Ala Pro Trp Glu Gln Gly Ser Arg Ala Ala Pro Leu Gln Glu
            340                 345                 350
Ala Val Pro Arg Thr Gln Ala Thr Gly Leu Thr Lys Pro Thr Leu Pro
        355                 360                 365
Pro Ser Pro Leu Met Ala Ala Arg Arg Arg Leu
    370                 375
```

<210> 26
<211> 1787
<212> DNA
<213> Homo sapiens

<400> 26

```
atgaggactc tccgcaggtt gaagttcatg agttcgccca gcctcagtga cctgggcaag 60
agagagccgg ccgccgccgc ggacgagcgg ggcacgcagc agcgccgggc ctgccgccaac 120
gccacctgga acagcatcca caacggggtg atcgccgtct tccagcgcaa ggggctgccc 180
gaccaggagc tcttcagcct caacgagggc gtccggcagc tgttgaagac agagctgggg 240
tcctttctca cggagtacct gcagaaccag ctgctgacaa aaggcatggt gatccttcgg 300
```

```
gacaagattc gcttctatga gggacagaag ctgctggact cactggcaga gacctgggac 360
ttcttcttca gtgacgtgct gcccatgctg caggccatct tctacccggt gcagggcaag 420
gagccatcgg tgcgccagct ggccctgctg cacttccgga atgccatcac cctcagtgtg 480
aagctagagg atgcgctggc ccgggcccat gcccgtgtgc ccccgtgcat cgtgcagatg 540
ctgctggtgc tgcaggcggt acatgagtcc aggggcgtga ctgaggacta cctgcgcctg 600
gagacgctgg tccagaaggt ggtgtcgcca tacctgggca cctacggcct ccaactccagc 660
gaggggccct tcacccattc ctgcatcctg gagctgcaga gagacaaggc ggcggcggct 720
gctgtgctgg gtgcagtgag gaagaggccc tcggtggtgc ccatggctgg ccaggatcct 780
gcgctgagca cgagtcaccc gttctacgac gtggccagac atggcattct gcaggtggca 840
ggggatgacc gctttggaag acgtgttgtc acgttcagct gctgccggat gccaccctcc 900
cacgagctgg accaccagcg gctgctggag tacaagaaga acttgaaggc cctctacgtg 960
gtgcacccca ccagcttcat caaggtcctg tggaacatct tgaagcccct catcagtcac 1020
aagtttggga agaaagtcat ctatttcaac tacctgagtg agctccacga acaccttaaa 1080
tacgaccagc tggtcatccc tcccgaagtt ttgcggtacg atgagaagct ccagagcctg 1140
cacgaggccc ggacgccgcc tcccaccaag acaccaccgc cgcggccccc gctgcccaca 1200
cagcagtttg gcgtcagtct gcaatacctc aaagacaaaa atcaaggcga actcatcccc 1260
cctgtgctga ggttcacagt gacgtacctg agagagaaag cctcccagag cacaactacg 1320
tcgtcctccg ctacctcatg ggcttcctgc atgcggtgtc ccgggagagc atcttcaaca 1380
aaatgaacag ctctaacctg gcctgtgtct tcgggctgaa tttgatctgg ccatcccagg 1440
gggtctcctc cctgagtgcc cttgtgcccc tgaacatgtt cactgaactg ctgatcgagt 1500
actatgaaaa gatcttcagc acccggagg cacctgggga gcacggcctg gcaccatggg 1560
aacaggggag caggtcagcc cctttgcagg aggctgtgcc acggacacaa gccacgggcc 1620
tcaccaaggc taccctacct ccgagtccca tgatggcagc cagaagacgt ctctagtgtt 1680
gcgaacactc tgtatatttc gagctacctc ccacacctgt ctgtgcactt gtatgttttg 1740
taaacttggc atctgtaaaa ataaccagcc attagatgaa ttcagaa                1787
```

<210> 27
<211> 461
<212> PRT
<213> Homo sapiens

<400> 27

```
Met Arg Thr Leu Arg Arg Leu Lys Phe Met Ser Ser Pro Ser Leu Ser
 1               5                  10                  15
Asp Leu Gly Lys Arg Glu Pro Ala Ala Ala Ala Asp Glu Arg Gly Thr
            20                  25                  30
Gln Gln Arg Arg Ala Cys Ala Asn Ala Thr Trp Asn Ser Ile His Asn
        35                  40                  45
Gly Val Ile Ala Val Phe Gln Arg Lys Gly Leu Pro Asp Gln Glu Leu
    50                  55                  60
Phe Ser Leu Asn Glu Gly Val Arg Gln Leu Leu Lys Thr Glu Leu Gly
65                  70                  75                  80
Ser Phe Phe Thr Glu Tyr Leu Gln Asn Gln Leu Leu Thr Lys Gly Met
                85                  90                  95
Val Ile Leu Arg Asp Lys Ile Arg Phe Tyr Glu Gly Gln Lys Leu Leu
            100                 105                 110
Asp Ser Leu Ala Glu Thr Trp Asp Phe Phe Phe Ser Asp Val Leu Pro
        115                 120                 125
Met Leu Gln Ala Ile Phe Tyr Pro Val Gln Gly Lys Glu Pro Ser Val
    130                 135                 140
Arg Gln Leu Ala Leu Leu His Phe Arg Asn Ala Ile Thr Leu Ser Val
145                 150                 155                 160
Lys Leu Glu Asp Ala Leu Ala Arg Ala His Ala Arg Val Pro Pro Ala
                165                 170                 175
Ile Val Gln Met Leu Leu Val Leu Gln Gly Val His Glu Ser Arg Gly
            180                 185                 190
Val Thr Glu Asp Tyr Leu Arg Leu Glu Thr Leu Val Gln Lys Val Val
        195                 200                 205
```

```
Ser Pro Tyr Leu Gly Thr Tyr Gly Leu His Ser Ser Glu Gly Pro Phe
    210                 215                 220
Thr His Ser Cys Ile Leu Glu Leu Gln Arg Asp Lys Ala Ala Ala Ala
225                 230                 235                 240
Ala Val Leu Gly Ala Val Arg Lys Arg Pro Ser Val Val Pro Met Ala
                245                 250                 255
Gly Gln Asp Pro Ala Leu Ser Thr Ser His Pro Phe Tyr Asp Val Ala
            260                 265                 270
Arg His Gly Ile Leu Gln Val Ala Gly Asp Asp Arg Phe Gly Arg Arg
        275                 280                 285
Val Val Thr Phe Ser Cys Cys Arg Met Pro Pro Ser His Glu Leu Asp
    290                 295                 300
His Gln Arg Leu Leu Glu Tyr Lys Lys Asn Leu Lys Ala Leu Tyr Val
305                 310                 315                 320
Val His Pro Thr Ser Phe Ile Lys Val Leu Trp Asn Ile Leu Lys Pro
                325                 330                 335
Leu Ile Ser His Lys Phe Gly Lys Lys Val Ile Tyr Phe Asn Tyr Leu
            340                 345                 350
Ser Glu Leu His Glu His Leu Lys Tyr Asp Gln Leu Val Ile Pro Pro
        355                 360                 365
Glu Val Leu Arg Tyr Asp Glu Lys Leu Gln Ser Leu His Glu Gly Arg
    370                 375                 380
Thr Pro Pro Pro Thr Lys Thr Pro Pro Pro Arg Pro Pro Leu Pro Thr
385                 390                 395                 400
Gln Gln Phe Gly Val Ser Leu Gln Tyr Leu Lys Asp Lys Asn Gln Gly
                405                 410                 415
Glu Leu Ile Pro Pro Val Leu Arg Phe Thr Val Thr Tyr Leu Arg Glu
            420                 425                 430
Lys Ala Ser Gln Ser Thr Thr Thr Ser Ser Ser Ala Thr Ser Trp Ala
        435                 440                 445
Ser Cys Met Arg Cys Pro Gly Arg Ala Ser Ser Thr Lys
    450                 455                 460
```

<210> 28
<211> 1176
<212> DNA
<213> Homo sapiens

<400> 28

```
acgtccgggg aggggccagg tgagcggcag acccggcacg caggtggggg ccggcggggt 60
ccgtggccag agctgcagag agacaaggcg gcggcggctg ctgtgctggg tgcagtgagg 120
aagaggccct cggtggtgcc catggctggc caggatcctg cgctgagcac gagtcacccg 180
ttctacgacg tggccagaca tggcattctg caggtggcag gggatgaccg ctttggaaga 240
cgtgttgtca cgttcagctg ctgccggatg ccaccctccc acgagctgga ccaccagcgg 300
ctgctggaca ggtacaagaa gaacttgaag gccctctacg tggtgcaccc caccagcttc 360
atcaaggtcc tgtggaacat cttgaagccc ctcatcagtc acaagtttgg gaagaaagtc 420
atctatttca actacctgag tgagctccac gaacacctta aatacgacca gctggtcatc 480
cctcccgaag ttttgcggta cgatgagaag ctccagagcc tgcacgaggg ccggacgccg 540
cctcccacca agacaccacc gccgcggccc ccgctgccca cacagcagtt tggcgtcagt 600
ctgcaatacc tcaaagacaa aaatcaaggc gaactcatcc cccctgtgct gaggttcaca 660
gtgacgtacc tgagagagaa aggcctccca gagcacaact acgtcgtcct ccgctacctc 720
atgggcttcc tgcatgcggt gtcccgggag agcatcttca acaaaatgaa cagctctaac 780
ctggcctgtg tcttcggggt gaatttgatc tggccatccc agggggtctc ctccctgagt 840
gcccttgtgc ccctgaacat gttcactgaa ctgctgatcg aqtactatga aaagatcttc 900
agcacccccg aggcacctgg ggagcacggc ctggcaccat gggaacaggg gagcagggca 960
gcccctttgc aggaggctgt gccacggaca caagccacgg gcctcaccaa gcctacccta 1020
cctccgagtc ccctgatggc agccagaaga cgtctctagt gttgcgaaca ctctgtatat 1080
```

ttcgagctac ctcccacacc tgtctgtgca cttgtatgtt ttgtaaactt ggcatctgta 1140
aaaataacca gccattagat gaattcagaa ccttct 1176

<210> 29
<211> 305
<212> PRT
<213> Homo sapiens

<400> 29
Met Ala Gly Gln Asp Pro Ala Leu Ser Thr Ser His Pro Phe Tyr Asp
1 5 10 15
Val Ala Arg His Gly Ile Leu Gln Val Ala Gly Asp Asp Arg Phe Gly
20 25 30
Arg Arg Val Val Thr Phe Ser Cys Cys Arg Met Pro Pro Ser His Glu
35 40 45
Leu Asp His Gln Arg Leu Leu Asp Arg Tyr Lys Lys Asn Leu Lys Ala
50 55 60
Leu Tyr Val Val His Pro Thr Ser Phe Ile Lys Val Leu Trp Asn Ile
65 70 75 80
Leu Lys Pro Leu Ile Ser His Lys Phe Gly Lys Lys Val Ile Tyr Phe
85 90 95
Asn Tyr Leu Ser Glu Leu His Glu His Leu Lys Tyr Asp Gln Leu Val
100 105 110
Ile Pro Pro Glu Val Leu Arg Tyr Asp Glu Lys Leu Gln Ser Leu His
115 120 125
Glu Gly Arg Thr Pro Pro Pro Thr Lys Thr Pro Pro Pro Arg Pro Pro
130 135 140
Leu Pro Thr Gln Gln Phe Gly Val Ser Leu Gln Tyr Leu Lys Asp Lys
145 150 155 160
Asn Gln Gly Glu Leu Ile Pro Pro Val Leu Arg Phe Thr Val Thr Tyr
165 170 175
Leu Arg Glu Lys Gly Leu Pro Glu His Asn Tyr Val Val Leu Arg Tyr
180 185 190
Leu Met Gly Phe Leu His Ala Val Ser Arg Glu Ser Ile Phe Asn Lys
195 200 205
Met Asn Ser Ser Asn Leu Ala Cys Val Phe Gly Leu Asn Leu Ile Trp
210 215 220
Pro Ser Gln Gly Val Ser Ser Leu Ser Ala Leu Val Pro Leu Asn Met
225 230 235 240
Phe Thr Glu Leu Leu Ile Glu Tyr Tyr Glu Lys Ile Phe Ser Thr Pro
245 250 255
Glu Ala Pro Gly Glu His Gly Leu Ala Pro Trp Glu Gln Gly Ser Arg
260 265 270
Ala Ala Pro Leu Gln Glu Ala Val Pro Arg Thr Gln Ala Thr Gly Leu
275 280 285
Thr Lys Pro Thr Leu Pro Pro Ser Pro Leu Met Ala Ala Arg Arg Arg
290 295 300
Leu
305

<210> 30
<211> 3257
<212> DNA
<213> Homo sapiens

<400> 30
atttccttct cccttcccc gccagcttcg catccatctc ccccaccccg taacccccte 60

ctgcctccat ccaccggggc tatggccgca gaagaggtat tgcagacggt ggaccattat 120
aagactgaga tagagaggct aaccaaggag ctcacggaga ccacccacga gaagatccag 180
gctgccgagt acgggctggt ggtgctggag gagaagctga ccctcaaaca gcagtatgat 240
gaactggagg ctgagtacga cagcctcaaa caggagctgg agcagctcaa agaggcattt 300
gggcagtcct tctccatcca ccggaaggtt gctgaagatg gagagactcg ggaggaaacg 360
cttctgcagg agtcagcatc gaaggaggct tactatctgg ggaagatctt ggagatgcag 420
aacgagctga aacagagccg ggctgtggtc actaatgtac aggcagaaaa cgagaggctc 480
accgcagtcg tgcaggatct gaaggagaac aatgagatgg tggagctaca gagaataccg 540
atgaaggatg aaatccgaga atataagttc cgggaggcac ggctccttca ggactatact 600
gaattggaag aagaaaatat cacattgcag aaactagtgt ccacgttgaa gcagaaccag 660
gttgaatacg aaggcttaaa gcatgagatt aagcgatttg aggaggagac ggtactgctg 720
aacagccagc tggaagatgc catccgattg aaagagattg ctgagcacca actggaagaa 780
gccctcgaga ctttaaaaaa tgaaagagag caaaagaaca acctgcggaa ggagctctcc 840
cagtatatca gcctcaatga taaccatatc agcatctcag tagatggact caaatttgcc 900
gaggatggga gtgaaccaaa caatgatgac aaaatgaacg gtcatatcca tgggcctctt 960
gtgaaactga atggagacta tcggactccc accttaagga aaggagagtc tctgaaccct 1020
gtctctgact tattcagtga gctgaacatt tcagaaatac agaagttgaa gcagcagctt 1080
atgcaggtag agcgggaaaa ggccattctt ttggccaacc tacaggagtc acagacacag 1140
ctggaacaca ccaaggggac actgacggag cagcatgagc gggtgcaccg gctcacagag 1200
cacgtcaatg ccatgagggg cctgcaaagc agcaaggagc tcaaggctga gctggacggg 1260
gagaagggcc gggactcagg ggaggaggcc catgactatg aggtggacat caatggttta 1320
gagatccttg aatgcaaata cagggtggca gtaactgagg tgattgatct gaagaactga 1380
attaaggcct taaaggagaa atataataaa tctgtagaaa actacactga tgagaaggcc 1440
aagtatgaga gtaaaatcca gatgtatgat gagcaggtga caagccttga gaagaccacc 1500
aaggagagtg gtgagaagat ggcccacatg gagaaggagt tgcaaaagat gaccagcata 1560
gccaacgaaa atcacagtac ccttaatacg gcccaggatg agttagtgac attcagtgag 1620
gagttagctc agctttacca ccatgtgtgt ctatgtaata atgaaactcc caacagggtc 1680
atgctggatt actatagggca gagcagagtc acccgcagtg gcagcctgaa agggcccgat 1740
gatcccagag gacttttgtc cccacgatta gccaggcggg gtgtgtcatc cccggtagaa 1800
acaaggacct cctctgaacc agttgcaaaa gaaagcacag agcccagcaa agaaccaagt 1860
ccaactaaga cccccacaat ctctcctgtt attactgccc caccgtcatc tccagtattg 1920
gatacaagtg acatccgcaa agagccaatg aatatctaca accttaatgc catcattccgg 1980
gaccaaatca agcatctgca gaaagctgtg gaccggtcct tgcaactgtc tcgtcaaaga 2040
gcagcagctc gggagctagc ccccatgatt gataaagaca aggaagcctt aatggaagag 2100
atcctcaagc taaagtccct gctgagcacc aaacgggagc agatcgccac attggagggcg 2160
gtgttgaaag ccaacaagca gacagctgag gtggcgctag ctaatctcaa gaacaaatat 2220
gaaaatgaaa aagcaatggt gactgaaacc atgacgaagc ttagaaatga actgaaggct 2280
ttgaaagaag atgctgcaac cttctcatcc ctgagaacaa tgtttgcaac aagatatgat 2340
gaatatgtca cccagttgga tgagatgcag agacagttag cagctgcaga ggatgagaag 2400
aagactctga acactttgtt acgaatggct atccagcaaa aactcgcctt gacccagagg 2460
ctggaggact tagagtttga ccatgagcag tcccgacgca gcaaaggcaa acttggaaag 2520
agcaagatcg cgagccctaa agtaagtggg gaggcatcag tcaccgtgcc caccatagac 2580
acttacctcc tgcatagtca gggcccacag acacccaaca ttcgggtcag cagtggcact 2640
cagaggaaag gacaattttc accttccctt tgtgatcaga gccgtcccag gatttcaggg 2700
gcttcctacc tacagaattt attaagagtt ccccctgatc ccacctccac agaatcattt 2760
cttctgaagg gccccccttc catgagtgaa ttcatccaag ggcaccggct cagcaaggaa 2820
aaaaggttaa ccgtggctcc accagattgt cagcagcctg ctgcctccgt accgccacag 2880
tgctcacaac tagccgggag gcaagactgc ccaactgtca gtcctgacac agtcctccct 2940
gaggagcagc cacattccag ctcccagtgc gcccctctcc actgtctctc caagcctcct 3000
caccccctagt cttcatctcc tgtggacgaa catctggggt ggaagttttg tagccacaca 3060
caggatactg cccaagatcc agcgggtgtt ttcttctcgg ttgttagatg tacaattgga 3120
ttaatgtcca tcgttttgga agacgagaaa gttgagaaga acacgaagca cagacccgta 3180
tgtgataaaa catttttgtgg tttctctgag tcacagataa acttctgcca tcaaatggct 3240
acagttcatt taaattt 3257

<210> 31
<211> 975
<212> PRT

<213> Homo sapiens

<400> 31

```
Met Ala Ala Glu Glu Val Leu Gln Thr Val Asp His Tyr Lys Thr Glu
 1               5                  10                  15
Ile Glu Arg Leu Thr Lys Glu Leu Thr Glu Thr Thr His Glu Lys Ile
            20                  25                  30
Gln Ala Ala Glu Tyr Gly Leu Val Val Leu Glu Glu Lys Leu Thr Leu
        35                  40                  45
Lys Gln Gln Tyr Asp Glu Leu Glu Ala Glu Tyr Asp Ser Leu Lys Gln
    50                  55                  60
Glu Leu Glu Gln Leu Lys Glu Ala Phe Gly Gln Ser Phe Ser Ile His
65                  70                  75                  80
Arg Lys Val Ala Glu Asp Gly Glu Thr Arg Glu Glu Thr Leu Leu Gln
                85                  90                  95
Glu Ser Ala Ser Lys Glu Ala Tyr Tyr Leu Gly Lys Ile Leu Glu Met
            100                 105                 110
Gln Asn Glu Leu Lys Gln Ser Arg Ala Val Val Thr Asn Val Gln Ala
        115                 120                 125
Glu Asn Glu Arg Leu Thr Ala Val Val Gln Asp Leu Lys Glu Asn Asn
    130                 135                 140
Glu Met Val Glu Leu Gln Arg Ile Arg Met Lys Asp Glu Ile Arg Glu
145                 150                 155                 160
Tyr Lys Phe Arg Glu Ala Arg Leu Leu Gln Asp Tyr Thr Glu Leu Glu
                165                 170                 175
Glu Glu Asn Ile Thr Leu Gln Lys Leu Val Ser Thr Leu Lys Gln Asn
            180                 185                 190
Gln Val Glu Tyr Glu Gly Leu Lys His Glu Ile Lys Arg Phe Glu Glu
        195                 200                 205
Glu Thr Val Leu Leu Asn Ser Gln Leu Glu Asp Ala Ile Arg Leu Lys
    210                 215                 220
Glu Ile Ala Glu His Gln Leu Glu Glu Ala Leu Glu Thr Leu Lys Asn
225                 230                 235                 240
Glu Arg Glu Gln Lys Asn Asn Leu Arg Lys Glu Leu Ser Gln Tyr Ile
                245                 250                 255
Ser Leu Asn Asp Asn His Ile Ser Ile Ser Val Asp Gly Leu Lys Phe
            260                 265                 270
Ala Glu Asp Gly Ser Glu Pro Asn Asn Asp Asp Lys Met Asn Gly His
        275                 280                 285
Ile His Gly Pro Leu Val Lys Leu Asn Gly Asp Tyr Arg Thr Pro Thr
    290                 295                 300
Leu Arg Lys Gly Glu Ser Leu Asn Pro Val Ser Asp Leu Phe Ser Glu
305                 310                 315                 320
Leu Asn Ile Ser Glu Ile Gln Lys Leu Lys Gln Gln Leu Met Gln Val
                325                 330                 335
Glu Arg Glu Lys Ala Ile Leu Leu Ala Asn Leu Gln Glu Ser Gln Thr
            340                 345                 350
Gln Leu Glu His Thr Lys Gly Ala Leu Thr Glu Gln His Glu Arg Val
        355                 360                 365
His Arg Leu Thr Glu His Val Asn Ala Met Arg Gly Leu Gln Ser Ser
    370                 375                 380
Lys Glu Leu Lys Ala Glu Leu Asp Gly Glu Lys Gly Arg Asp Ser Gly
385                 390                 395                 400
Glu Glu Ala His Asp Tyr Glu Val Asp Ile Asn Gly Leu Glu Ile Leu
                405                 410                 415
Glu Cys Lys Tyr Arg Val Ala Val Thr Glu Val Ile Asp Leu Lys Ala
            420                 425                 430
Glu Ile Lys Ala Leu Lys Glu Lys Tyr Asn Lys Ser Val Glu Asn Tyr
```

```
        435                 440                 445
Thr Asp Glu Lys Ala Lys Tyr Glu Ser Lys Ile Gln Met Tyr Asp Glu
    450                 455                 460
Gln Val Thr Ser Leu Glu Lys Thr Thr Lys Glu Ser Gly Glu Lys Met
465                 470                 475                 480
Ala His Met Glu Lys Glu Leu Gln Lys Met Thr Ser Ile Ala Asn Glu
                485                 490                 495
Asn His Ser Thr Leu Asn Thr Ala Gln Asp Glu Leu Val Thr Phe Ser
            500                 505                 510
Glu Glu Leu Ala Gln Leu Tyr His His Val Cys Leu Cys Asn Asn Glu
        515                 520                 525
Thr Pro Asn Arg Val Met Leu Asp Tyr Tyr Arg Gln Ser Arg Val Thr
    530                 535                 540
Arg Ser Gly Ser Leu Lys Gly Pro Asp Asp Pro Arg Gly Leu Leu Ser
545                 550                 555                 560
Pro Arg Leu Ala Arg Arg Gly Val Ser Ser Pro Val Glu Thr Arg Thr
                565                 570                 575
Ser Ser Glu Pro Val Ala Lys Glu Ser Thr Glu Pro Ser Lys Glu Pro
            580                 585                 590
Ser Pro Thr Lys Thr Pro Thr Ile Ser Pro Val Ile Thr Ala Pro Pro
        595                 600                 605
Ser Ser Pro Val Leu Asp Thr Ser Asp Ile Arg Lys Glu Pro Met Asn
    610                 615                 620
Ile Tyr Asn Leu Asn Ala Ile Ile Arg Asp Gln Ile Lys His Leu Gln
625                 630                 635                 640
Lys Ala Val Asp Arg Ser Leu Gln Leu Ser Arg Gln Arg Ala Ala Ala
                645                 650                 655
Arg Glu Leu Ala Pro Met Ile Asp Lys Asp Lys Glu Ala Leu Met Glu
            660                 665                 670
Glu Ile Leu Lys Leu Lys Ser Leu Leu Ser Thr Lys Arg Glu Gln Ile
        675                 680                 685
Ala Thr Leu Arg Ala Val Leu Lys Ala Asn Lys Gln Thr Ala Glu Val
    690                 695                 700
Ala Leu Ala Asn Leu Lys Asn Lys Tyr Glu Asn Glu Lys Ala Met Val
705                 710                 715                 720
Thr Glu Thr Met Thr Lys Leu Arg Asn Glu Leu Lys Ala Leu Lys Glu
                725                 730                 735
Asp Ala Ala Thr Phe Ser Ser Leu Arg Thr Met Phe Ala Thr Arg Cys
            740                 745                 750
Asp Glu Tyr Val Thr Gln Leu Asp Glu Met Gln Arg Gln Leu Ala Ala
        755                 760                 765
Ala Glu Asp Glu Lys Lys Thr Leu Asn Thr Leu Leu Arg Met Ala Ile
    770                 775                 780
Gln Gln Lys Leu Ala Leu Thr Gln Arg Leu Glu Asp Leu Glu Phe Asp
785                 790                 795                 800
His Glu Gln Ser Arg Arg Ser Lys Gly Lys Leu Gly Lys Ser Lys Ile
                805                 810                 815
Gly Ser Pro Lys Val Ser Gly Glu Ala Ser Val Thr Val Pro Thr Ile
            820                 825                 830
Asp Thr Tyr Leu Leu His Ser Gln Gly Pro Gln Thr Pro Asn Ile Arg
        835                 840                 845
Val Ser Ser Gly Thr Gln Arg Lys Arg Gln Phe Ser Pro Ser Leu Cys
    850                 855                 860
Asp Gln Ser Arg Pro Arg Thr Ser Gly Ala Ser Tyr Leu Gln Asn Leu
865                 870                 875                 880
Leu Arg Val Pro Pro Asp Pro Thr Ser Thr Glu Ser Phe Leu Leu Lys
                885                 890                 895
Gly Pro Pro Ser Met Ser Glu Phe Ile Gln Gly His Arg Leu Ser Lys
```

```
            900                 905                 910
Glu Lys Arg Leu Thr Val Ala Pro Pro Asp Cys Gln Gln Pro Ala Ala
        915                 920                 925
Ser Val Pro Pro Gln Cys Ser Gln Leu Ala Gly Arg Gln Asp Cys Pro
    930                 935                 940
Thr Val Ser Pro Asp Thr Ala Leu Pro Glu Glu Gln Pro His Ser Ser
945                 950                 955                 960
Ser Gln Cys Ala Pro Leu His Cys Leu Ser Lys Pro Pro His Pro
                965                 970                 975
```

<210> 32
<211> 2717
<212> DNA
<213> Homo sapiens

<400> 32

```
cagggtaacg ctgtcttgtg gacccgcact tcccacccga gacctctcac tgagcccgag 60
ccgcgcgcga catgagccac gggaaggaa ccgacatgct cccggagatc gccgccgccg 120
tgggcttcct ctccagcctc ctgaggaccc ggggctgcgt gagcgagcag atggcttaagg 180
tcttcagcgg ggcgctccag gaggcactca cagagcacta caaaacaccac tggtttcccg 240
aaaagccgtc caaggggctcc ggctaccgct gcattcgcat caacccacaag atggacccca 300
tcatcagcag ggtggccagc cagatcggac tcagccagcc ccagctgcac caqctgctgc 360
ccagcgagct gaccctgtgg gtggacccct atgaggtgtc ctaccgcatt ggggaggacg 420
gctccatctg cgtcttgtac gaggaggccc cactggccgc ctcctgtggg ctcctcacct 480
gcaagaacca agtgctgctg ggccggaqca gcccctccaa gaactacgtg atqqcagtct 540
ccagctaggc ccttccgccc ccgccctggg cgccgccgtg ctcatgctgc cgtgacaaca 600
ggccaccaca tacctcaacc tggggaactg tatttttaaa tgaagagcta tttatatata 660
ttatttttt ttaagaaagg aggaaaaqaa accaaaagtt ttttttaaga aaaaaaatcc 720
ttcaagggag ctgcttggaa gtggcctccc caggtgcctt tggagagaac tgttgcgtgc 780
ttgagtctgt gagccagtgt ctgcctatag gaggggagc tgttagggg tagacctagc 840
caaggagaag tggagacgt ttggctagca ccccaggaag atgtgagagg gagcaagcaa 900
ggttagcaac tgtgaacaga gaggtcggga tttgccctgg gggaggaaga gaggccaagt 960
tcagagctct ctgtctcccc caqccaqaca cctgcatccc tggctcctct attactcagg 1020
ggcattcatg cctggactta aacaatacta tgttatcttt tcttttattt ttctaatgag 1080
gtcctgggca gagagtgaaa aggcctctcc tgattcctac tgtcctaagc tgcttttctt 1140
gaaatcatga cttgtttcta attctaccct caggggcctg tagatgttgc tttccagcca 1200
ggaatctaaa gctttgggtt ttctgagggg gggaggaggg aactgaggtt tatggggtt 1260
aggatggaag ggaactctgc acaaaacctt tgcttttgcta gtgctgcttt gtgtgtatgt 1320
gtggcaaata atttgggggt gatttgcaat gaaatttttg gacccaaaga gtatccactg 1380
gggatgtttt ttggccaaaa ctcttcctttt tggaaccaca tgaaagtctt gatgctgctg 1440
ccatgatccc tttgagaggt ggctcaaaag ctacaaggga ctccaggtcc tttattactg 1500
ccttcttttc aaaagcacaa ctctcctcta accctccact ccccttccc ttctggtcgg 1560
gtcatagagc tacogtattt tctaggacaa gagttctcag tcactgtgca atatgccccc 1620
tggqtcccag gaggtctgg aggaaaactg gctatcagaa cctcctgatg ccctggtggg 1680
cttagggaac catctctcct gctctccttg ggatgatggc tggctagtca gccttgcatg 1740
tattcttgg ctgaatggga gagtgcccca tgttctgcaa gactacttgg tattcttgta 1800
gggccgacac taaataaaag ccaaaccttg ggcactgttt tttctccctg gtgctcagag 1860
cacctgtggg aaaggttgct gtctgtctca gtacaatcca aatttgtcgt agacttgtgc 1920
aatatatact gttgtgggtt ggagaaaagt ggaaagctac actgggaaga aactcccttc 1980
cttcaattc tcagtgacat tgatgagggg tcctcaaaag acctcgagtt tcccaaaccg 2040
aatcacctta agaaggacag ggctagggca tttggccagg atggccaccc tcctgctgtt 2100
gcccctagt gaggaatctt caccccactt cctctacccc caggttctcc tccccacagc 2160
cagtcccctt tcctggattt ctaaactgct caattttgac tcaaaggtgc tatttaccaa 2220
acactctccc taccattcc tgccagctct gcctcctttt caactctcca cattttgtat 2280
tgccttccca gacctgcttc cagtctttat tgcttaaag ttcactttgg gcccacagac 2340
ccaagagcta attttctggt ttgtgggttg aaacaaagct gtgaatcact gcaggctgtg 2400
ttcttgcatc ttgtctgcaa acaggtccct gccttttttag aagcagcctc atggtctcat 2460
```

gcttaatctt gtctctcttc tcttctttat gatgttcact ttaaaaacaa caaaacccct 2520
gagctggact gttgagcagg cctgtctctc ctattaagta aaaataaata gtagtagtat 2580
gtttgtaagc tattctgaca gaaaagacaa aggttactaa ttgtatgata gtgtttttat 2640
atggaagaat gtacagctta tggacaaatg tacacctttt tgttacttta ataaaaatgt 2700
agtaggataa aaaaaaa 2717

<210> 33
<211> 158
<212> PRT
<213> Homo sapiens

<400> 33
Met Ser His Gly Lys Gly Thr Asp Met Leu Pro Glu Ile Ala Ala Ala
1 5 10 15
Val Gly Phe Leu Ser Ser Leu Leu Arg Thr Arg Gly Cys Val Ser Glu
20 25 30
Gln Arg Leu Lys Val Phe Ser Gly Ala Leu Gln Glu Ala Leu Thr Glu
35 40 45
His Tyr Lys His His Trp Phe Pro Glu Lys Pro Ser Lys Gly Ser Gly
50 55 60
Tyr Arg Cys Ile Arg Ile Asn His Lys Met Asp Pro Ile Ile Ser Arg
65 70 75 80
Val Ala Ser Gln Ile Gly Leu Ser Gln Pro Gln Leu His Gln Leu Leu
85 90 95
Pro Ser Glu Leu Thr Leu Trp Val Asp Pro Tyr Glu Val Ser Tyr Arg
100 105 110
Ile Gly Glu Asp Gly Ser Ile Cys Val Leu Tyr Glu Glu Ala Pro Leu
115 120 125
Ala Ala Ser Cys Gly Leu Leu Thr Cys Lys Asn Gln Val Leu Leu Gly
130 135 140
Arg Ser Ser Pro Ser Lys Asn Tyr Val Met Ala Val Ser Ser
145 150 155

<210> 34
<211> 5471
<212> DNA
<213> Homo sapiens

<400> 34
acagcttgca acactcggca tcttttctgg aggcgcctcc ttcagcagcc gcagatggca 60
tccggctgcg ggctcggggc tcgcaattga ttctcccctt tgcccacctc gagtccacgg 120
acgcacctct cccttcccct cctcccttcg cgcttctggg tctgagccca gctcgcgacc 180
gccgggcaga ggatcagtcg cggcggccga ggctgagcag cagcgctctc gctcctgac 240
ctggggagaa gcgcccaccc gggagagctg atccccggct gcctccagcg ccccccacct 300
tttgcactcc aagccgggcg ctccagagac cccgctcccc aggcgccact atgctggacc 360
cttcgtccag cgaagaagaa tcggatgaga tcgtggagga ggagagcggc aaggaggtgc 420
tcggctcggc cccgtccggc gcgcgcctgt ctcccagccg taccagcgag ggctcggccg 480
gcagcgccgg gctgggggcc ggcggccgcg gcgccggagc cgggtgggct gcaggcggcg 540
gcgggggcag cggcgcgagc agcggccggcg gggccgggcg gctgcaaccc agcagccgcg 600
ctggcggcgg ccggcgcctcc agccccagcc cgtcggtggt gagcgagaag gagaaggaag 660
agttggagcg gctgcagaaa gaggaggagg agaggaagaa gaggctgcag ctgtatgtgt 720
tcgtgatgcg ctgcatcgcc taccccttta atgccaagca gcccaccgac atggctcgcc 780
ggcagcagaa gatcagcaaa cagcagctgc agacagtcaa ggaccggttt caggctttcc 840
tcaatgggga aacccagatc atggctgacg aagccttcat gaacgctgtg cagagttact 900
atgaggtgtt cctgaagagc gaccgtgtgg cccgcatggt tcagagtgga ggctgttccg 960
ccaacgactc ccgggaggtc ttcaagaagc acattgagaa gagagtggcc agcctgcctg 1020
agattgacgg cctcagcaag gagactgtgc tgagctcctg gatggccaaa tttgatgcca 1080

tctaccgtgg agaagaggac ccgcggaagc agcaggcccg gatgacagcc agcgcagcct 1140
ccgagctgat tctgagcaag gagcaactct atgagatgtt ccagaacatt cttggatcca 1200
agaagttcga acatcagctc ctttacaatg cctgccagct ggacaatcca gatgagcaag 1260
cagcccagat cagacgagag ctggatggac gtctacaaat ggcagaccaa atagccaggg 1320
aacgcaaatt tcccaagttt gtatccaaag aaatggaaaa catgtacatt gaggagctga 1380
agtcatctgt caacctgctc atggccaact tggagagcat gccggtatcc aaggcgggg 1440
agttcaagct ccagaaactc aaacgcagcc acaatgcttc catcatcgac atgggcgagg 1500
agagtgagaa ccagtctcac aagtcagatg tcgtgctgtc tttctcattg gaggtggtaa 1560
ttatggaagt ccaaggcctc aaatctttgg ctccaaatcg catcgtatat tgcacaatgg 1620
aggtggaagg aggagagaaa ctacagactg atcaggccga ggcttctaaa ccaacctggg 1680
gcacccaggg tgacttctcc acaacccatg cactgccagc tgtgaaggtg aagctgttca 1740
cagagagcac aggcgtcctg gcgttggagg acaaggagct tgggcgggtt attctccatc 1800
ccaccccgaa cagccccaaa cagtcagagt ggcacaaaat gacagtctcc aaaaactgcc 1860
ccgaccaaga tctcaaaatc aaacttgctg tccgaatgga taagcctcaa aacatgaagc 1920
attctgggta tttatgggcc atcggtaaga atgtctggaa gagatggaag aaaaggtttt 1980
ttgtattggt gcaggtcagt cagtacacgt ttgccatgtg cagttatccg gagagaaaag 2040
cggagcctca ggaacttcta caattggatg gctacactgt ggattacacc gacccccagc 2100
caggtttgga gggtggccga gccttcttca atgctgtcaa ggagggagac accgtgatat 2160
ttgccagtga cgatgaacaa gaccgcatcc tgtgggtcca ggccatgtat cgggccacgg 2220
ggcagtcaca caagcctgtg cccccgaccc aagtccagaa actcaacgcc aagggaggaa 2280
atgtacctca gctggatgcc cctatctctc aattttacgc agatagagct caaaaacatg 2340
gcatggatga attatctctt tccaacccct gtaacttgga ccacgcttcc ctctttgaga 2400
tggtacaacg ccttactttg gatcacagac ttaatgattc ctattcttgc ctgggctggt 2460
tcagtcctgg ccaggtgttt gtactagacg agtattgcgc ccgaaatgga gtccgggggt 2520
gtcaccgaca tctctgctac ctcagagact tgcttgaacg ggcagaaaat ggcgccatga 2580
tcgaccccac cctcttctac tacagctttg ccttctgtgc atcccatgtc catcgtgaaca 2640
ggcctgatgg aattggaact gtgactgttg aagaaaagga acgttttgaa gaaatcaaag 2700
agaggctccg agttctgcta gaaaatcaga ttacacattt taggtattgc tttccatttg 2760
gtcgacctga aggtgctttg aaagctactc tctcactctt ggaaaggggtt ttgatgaaag 2820
atattgttac cccagtgcca caagaggagg taaaaacagt tatccgtaaa tgtctggaac 2880
aggctgcgtt agtcaactat tctcggctct cagagtatgc caaaatcgaa gagaatcaaa 2940
aggatgcaga aaatgtaggc cggttaatca ctcctgccaa aaagcttgaa gatacaatac 3000
gtcttgctga actagtcatt gaagttcttc agcaaaatga ggagcaccac gcagagccac 3060
atgttgataa aggagaagcc tttgcgtggt ggtcagattt aatggtggag catgcggaga 3120
cgttcctgtc actctttgca gtagacatgg atgcagcctt agaggtgcaa cctccagaca 3180
catgggacag ttttccacta tttcagctgc tgaatgattt tctccgtact gactataatt 3240
tgtgcaatgg aaaatttcac aaacacctgc aagacctgtt tgccccactt gttgttagat 3300
atgtggatct gatggagtcc tcaattgcac aatccattca caggggcttt gagcgggagt 3360
catgggaacc agtcaagagt ttaaccagta acctacccaa tgtgaaccta cccaatgtga 3420
accttcccaa agtaccaaat ctaccagtta acatccctct aggcatccca caaatgccta 3480
ctttttcggc accgtcatgg atggctgcta tatatgatgc ggataatggg tcaggcacct 3540
cagaagatct gttttggaaa cttgacgccc ttcagacctt cattcgggac ctgcactggc 3600
ctgaagaaga gtttggaaag cacctggaac aacggctgaa gttgatggca agtgacatga 3660
tcgaatcttg tgtcaaaaga accaggattg catttgaagt taagctgcaa aaaaccagtc 3720
gatcaacaga ttttcgagtc ccacagtcaa tatgcaccat gtttaatgtt atggttgatg 3780
ccaaagctca atcaacaaaa ctttgcagca tggaaatggg ccaagagcat caatacctt 3840
caaaaataga cgaactaatt gaagaaactg ttaaagaaat gataacactc ttggttgcaa 3900
agttcgttac tatcttggaa ggagtgctgg caaaattatc cagatatgac gaagggactt 3960
tgttttcttc ttttctgtca tttaccgtga aggcagcttc caaatatgtg gatgtaccta 4020
aacccgggat ggacgtggcc gacgcctacg tgacttttcg ccgccattct caggatgtcc 4080
tgcgtgataa ggtcaatgag gagatgtaca tagaaggtt atttgatcaa tggtacaaca 4140
gctccatgaa cgtgatctgc acctggttga cggaccggat ggacttacag cttcatattt 4200
atcagttgaa aacactaatt agggtggtaa agaaaaccta cagagatttc cgattgcaag 4260
gggtcctgga ctccacctta aacagcaaga cctatgaaac gatccggaac cgtctcactg 4320
tggaggaagc cacagcatca gtgagtgaag gtgggggact gcagggcatc agcatgaagg 4380
acagcgatga ggaagacgaa gaagacgatt agaccatttg gtcctagagt ctgctgggac 4440
agagtcctgt aatcagtgca tgtccttagt ctgttagtta aacccattag gaatttttctg 4500
tcaactacca tgcccatgag atgtttatca atacaactgc catttagct atgtggtacc 4560

aagattagca aatgaccttc atatccactg atttcctgat gtccatgtct atatgtttac 4620
aagcaatatg gagcaccatt ctttaaatac tgttcatgga gaatacatag tctaaccact 4680
aggcgtgtcc ctgttatcag caaagatcaa tgatgcttca ttcatgtact atgtatgcat 4740
tggtggtaaa tggatgtgag ggcaagtaca tcaagtacat tcactctgtt tcacgtatgt 4800
ggatgccagt taattaaatg agtacgtaaa taaattaatt aaaacacata gatctgcttt 4860
gtgtttttat ttttattttt tgaaaaacaa aaggcaagtc tccaacaatt aacttttgat 4920
gctttctgtt cccctaaaac caaaaaatga acccctgtg tcgttgttaa cccatccttt 4980
catttactca tataattagc caaaaaaaaa aggatggcta cataccaatg gattgattct 5040
cttaattgcc acggcaaggg ggcgatccta tcatgactta acatcaagcg cgcagttcaa 5100
aactactgtc ttctgtcaaa gttttctcct cttaaatgtt atttgctttt tacgtctcaa 5160
ctgtgtatgt aaaaaaaacg aatatttaaa ttacaacccct agactaaaaa tgtgtttata 5220
ataagatgtg gatatttcct tcagtagatt gtaaccataa tttaaattat tttgttccac 5280
actgtttttt atatctgtca tgtacattgc attttgatct gtaactgcac aaccctgggg 5340
tttgctgcag agctatttct ttccatgtaa agtagtggat ccatcttgct tttgccttat 5400
ataaagccta cagttatgga agtgtggaaa actgtggctt ctcaataaat attcagatgt 5460
cctaagaata t 5471

<210> 35
<211> 1390
<212> PRT
<213> Homo sapiens

<400> 35
Pro Gly Glu Lys Arg Pro Pro Gly Arg Ala Asp Pro Arg Leu Pro Pro
1 5 10 15
Ala Pro Pro Thr Phe Cys Thr Pro Ser Arg Gly Leu Gln Arg Pro Arg
20 25 30
Ser Pro Gly Ala Thr Met Leu Asp Pro Ser Ser Ser Glu Glu Glu Ser
35 40 45
Asp Glu Ile Val Glu Glu Glu Ser Gly Lys Glu Val Leu Gly Ser Ala
50 55 60
Pro Ser Gly Ala Arg Leu Ser Pro Ser Arg Thr Ser Glu Gly Ser Ala
65 70 75 80
Gly Ser Ala Gly Leu Gly Gly Gly Gly Ala Gly Ala Gly Ala Gly Val
85 90 95
Gly Ala Gly Gly Gly Gly Ser Gly Ala Ser Ser Gly Gly Gly Ala
100 105 110
Gly Gly Leu Gln Pro Ser Ser Arg Ala Gly Gly Gly Arg Pro Ser Ser
115 120 125
Pro Ser Pro Ser Val Val Ser Glu Lys Glu Lys Glu Leu Glu Arg
130 135 140
Leu Gln Lys Glu Glu Glu Glu Arg Lys Lys Arg Leu Gln Leu Tyr Val
145 150 155 160
Phe Val Met Arg Cys Ile Ala Tyr Pro Phe Asn Ala Lys Gln Pro Thr
165 170 175
Asp Met Ala Arg Arg Gln Gln Lys Ile Ser Lys Gln Gln Leu Gln Thr
180 185 190
Val Lys Asp Arg Phe Gln Ala Phe Leu Asn Gly Glu Thr Gln Ile Met
195 200 205
Ala Asp Glu Ala Phe Met Asn Ala Val Gln Ser Tyr Tyr Glu Val Phe
210 215 220
Leu Lys Ser Asp Arg Val Ala Arg Met Val Gln Ser Gly Gly Cys Ser
225 230 235 240
Ala Asn Asp Ser Arg Glu Val Phe Lys Lys His Ile Glu Lys Arg Val
245 250 255
Arg Ser Leu Pro Glu Ile Asp Gly Leu Ser Lys Glu Thr Val Leu Ser
260 265 270
Ser Trp Met Ala Lys Phe Asp Ala Ile Tyr Arg Gly Glu Glu Asp Pro

```
        275                 280                 285
Arg Lys Gln Gln Ala Arg Met Thr Ala Ser Ala Ala Ser Glu Leu Ile
    290                 295                 300
Leu Ser Lys Glu Gln Leu Tyr Glu Met Phe Gln Asn Ile Leu Gly Ile
305                 310                 315                 320
Lys Lys Phe Glu His Gln Leu Leu Tyr Asn Ala Cys Gln Leu Asp Asn
                325                 330                 335
Pro Asp Glu Gln Ala Ala Gln Ile Arg Arg Glu Leu Asp Gly Arg Leu
            340                 345                 350
Gln Met Ala Asp Gln Ile Ala Arg Glu Arg Lys Phe Pro Lys Phe Val
        355                 360                 365
Ser Lys Glu Met Glu Asn Met Tyr Ile Glu Glu Leu Lys Ser Ser Val
    370                 375                 380
Asn Leu Leu Met Ala Asn Leu Glu Ser Met Pro Val Ser Lys Gly Gly
385                 390                 395                 400
Glu Phe Lys Leu Gln Lys Leu Lys Arg Ser His Asn Ala Ser Ile Ile
                405                 410                 415
Asp Met Gly Glu Glu Ser Glu Asn Gln Leu Ser Lys Ser Asp Val Val
            420                 425                 430
Leu Ser Phe Ser Leu Glu Val Val Ile Met Glu Val Gln Gly Leu Lys
        435                 440                 445
Ser Leu Ala Pro Asn Arg Ile Val Tyr Cys Thr Met Glu Val Glu Gly
    450                 455                 460
Gly Glu Lys Leu Gln Thr Asp Gln Ala Glu Ala Ser Lys Pro Thr Trp
465                 470                 475                 480
Gly Thr Gln Gly Asp Phe Ser Thr Thr His Ala Leu Pro Ala Val Lys
                485                 490                 495
Val Lys Leu Phe Thr Glu Ser Thr Gly Val Leu Ala Leu Glu Asp Lys
            500                 505                 510
Glu Leu Gly Arg Val Ile Leu His Pro Thr Pro Asn Ser Pro Lys Gln
        515                 520                 525
Ser Glu Trp His Lys Met Thr Val Ser Lys Asn Cys Pro Asp Gln Asp
    530                 535                 540
Leu Lys Ile Lys Leu Ala Val Arg Met Asp Lys Pro Gln Asn Met Lys
545                 550                 555                 560
His Ser Gly Tyr Leu Trp Ala Ile Gly Lys Asn Val Trp Lys Arg Trp
                565                 570                 575
Lys Lys Arg Phe Phe Val Leu Val Gln Val Ser Gln Tyr Thr Phe Ala
            580                 585                 590
Met Cys Ser Tyr Arg Glu Lys Lys Ala Glu Pro Gln Glu Leu Leu Gln
        595                 600                 605
Leu Asp Gly Tyr Thr Val Asp Tyr Thr Asp Pro Gln Pro Gly Leu Glu
    610                 615                 620
Gly Gly Arg Ala Phe Phe Asn Ala Val Lys Glu Gly Asp Thr Val Ile
625                 630                 635                 640
Phe Ala Ser Asp Asp Glu Gln Asp Arg Ile Leu Trp Val Gln Ala Met
                645                 650                 655
Tyr Arg Ala Thr Gly Gln Ser His Lys Pro Val Pro Pro Thr Gln Val
            660                 665                 670
Gln Lys Leu Asn Ala Lys Gly Gly Asn Val Pro Gln Leu Asp Ala Pro
        675                 680                 685
Ile Ser Gln Phe Tyr Ala Asp Arg Ala Gln Lys His Gly Met Asp Glu
    690                 695                 700
Phe Ile Ser Ser Asn Pro Cys Asn Phe Asp His Ala Ser Leu Phe Glu
705                 710                 715                 720
Met Val Gln Arg Leu Thr Leu Asp His Arg Leu Asn Asp Ser Tyr Ser
                725                 730                 735
Cys Leu Gly Trp Phe Ser Pro Gly Gln Val Phe Val Leu Asp Glu Tyr
```

```
            740                 745                 750
Cys Ala Arg Asn Gly Val Arg Gly Cys His Arg His Leu Cys Tyr Leu
        755                 760                 765
Arg Asp Leu Leu Glu Arg Ala Glu Asn Gly Ala Met Ile Asp Pro Thr
    770                 775                 780
Leu Leu His Tyr Ser Phe Ala Phe Cys Ala Ser His Val His Gly Asn
785                 790                 795                 800
Arg Pro Asp Gly Ile Gly Thr Val Thr Val Glu Glu Lys Glu Arg Phe
                805                 810                 815
Glu Glu Ile Lys Glu Arg Leu Arg Val Leu Leu Glu Asn Gln Ile Thr
            820                 825                 830
His Phe Arg Tyr Cys Phe Pro Phe Gly Arg Pro Glu Gly Ala Leu Lys
        835                 840                 845
Ala Thr Leu Ser Leu Leu Glu Arg Val Leu Met Lys Asp Ile Val Thr
    850                 855                 860
Pro Val Pro Gln Glu Glu Val Lys Thr Val Ile Arg Lys Cys Leu Glu
865                 870                 875                 880
Gln Ala Ala Leu Val Asn Tyr Ser Arg Leu Ser Glu Tyr Ala Lys Ile
                885                 890                 895
Glu Glu Asn Gln Lys Asp Ala Glu Asn Val Gly Arg Leu Ile Thr Pro
            900                 905                 910
Ala Lys Lys Leu Glu Asp Thr Ile Arg Leu Ala Glu Leu Val Ile Glu
        915                 920                 925
Val Leu Gln Gln Asn Glu Glu His His Ala Glu Pro His Val Asp Lys
    930                 935                 940
Gly Glu Ala Phe Ala Trp Trp Ser Asp Leu Met Val Glu His Ala Glu
945                 950                 955                 960
Thr Phe Leu Ser Leu Phe Ala Val Asp Met Asp Ala Ala Leu Glu Val
                965                 970                 975
Gln Pro Pro Asp Thr Trp Asp Ser Phe Pro Leu Phe Gln Leu Leu Asn
            980                 985                 990
Asp Phe Leu Arg Thr Asp Tyr Asn Leu Cys Asn Gly Lys Phe His Lys
        995                 1000                1005
His Leu Gln Asp Leu Phe Ala Pro Leu Val Val Arg Tyr Val Asp Leu
    1010                1015                1020
Met Glu Ser Ser Ile Ala Gln Ser Ile His Arg Gly Phe Glu Arg Glu
1025                1030                1035                1040
Ser Trp Glu Pro Val Lys Ser Leu Thr Ser Asn Leu Pro Asn Val Asn
                1045                1050                1055
Leu Pro Asn Val Asn Leu Pro Lys Val Pro Asn Leu Pro Val Asn Ile
            1060                1065                1070
Pro Leu Gly Ile Pro Gln Met Pro Thr Phe Ser Ala Pro Ser Trp Met
        1075                1080                1085
Ala Ala Ile Tyr Asp Ala Asp Asn Gly Ser Gly Thr Ser Glu Asp Leu
    1090                1095                1100
Phe Trp Lys Leu Asp Ala Leu Gln Thr Phe Ile Arg Asp Leu His Trp
1105                1110                1115                1120
Pro Glu Glu Glu Phe Gly Lys His Leu Glu Gln Arg Leu Lys Leu Met
                1125                1130                1135
Ala Ser Asp Met Ile Glu Ser Cys Val Lys Arg Thr Arg Ile Ala Phe
            1140                1145                1150
Glu Val Lys Leu Gln Lys Thr Ser Arg Ser Thr Asp Phe Arg Val Pro
        1155                1160                1165
Gln Ser Ile Cys Thr Met Phe Asn Val Met Val Asp Ala Lys Ala Gln
    1170                1175                1180
Ser Thr Lys Leu Cys Ser Met Glu Met Gly Gln Glu His Gln Tyr His
1185                1190                1195                1200
Ser Lys Ile Asp Glu Leu Ile Glu Glu Thr Val Lys Glu Met Ile Thr
```

```
                1205                1210                1215
Leu Leu Val Ala Lys Phe Val Thr Ile Leu Glu Gly Val Leu Ala Lys
            1220                1225                1230
Leu Ser Arg Tyr Asp Glu Gly Thr Leu Phe Ser Ser Phe Leu Ser Phe
        1235                1240                1245
Thr Val Lys Ala Ala Ser Lys Tyr Val Asp Val Pro Lys Pro Gly Met
    1250                1255                1260
Asp Val Ala Asp Ala Tyr Val Thr Phe Val Arg His Ser Gln Asp Val
1265                1270                1275                1280
Leu Arg Asp Lys Val Asn Glu Glu Met Tyr Ile Glu Arg Leu Phe Asp
                1285                1290                1295
Gln Trp Tyr Asn Ser Ser Met Asn Val Ile Cys Thr Trp Leu Thr Asp
            1300                1305                1310
Arg Met Asp Leu Gln Leu His Ile Tyr Gln Leu Lys Thr Leu Ile Arg
        1315                1320                1325
Val Val Lys Lys Thr Tyr Arg Asp Phe Arg Leu Gln Gly Val Leu Asp
    1330                1335                1340
Ser Thr Leu Asn Ser Lys Thr Tyr Glu Thr Ile Arg Asn Arg Leu Thr
1345                1350                1355                1360
Val Glu Glu Ala Thr Ala Ser Val Ser Glu Gly Gly Gly Leu Gln Gly
                1365                1370                1375
Ile Ser Met Lys Asp Ser Asp Glu Glu Asp Glu Asp Asp
            1380                1385                1390
```

<210> 36
<211> 4828
<212> DNA
<213> Homo sapiens

<400> 36

```
agtggcgtcg gaactgcaaa gcacctgtga gcttgcggaa gtcagttcag actccagccc 60
gctccagccc ggcccgaccc gaccgcaccc ggcgcctgcc ctcgctcggc gtccccggcc 120
agccatgggc ccttggagcc gcagcctctc ggcgctgctg ctgctgctgc aggtctcctc 180
ttggctctgc caggagccgg agccctgcca cccctggctt gacgccgaga gctacacgtt 240
cacggtgccc cggcgccacc tggagagagg ccgcgtcctg ggcagagtga attttgaaga 300
ttgcaccggt cgacaaagga cagcctattt ttcctctgac acccgattca aagtgggcac 360
agatggtgtg attacagtca aaaggcctct acggtttcat aaccccacga tccatttctt 420
ggtctacgcc tgggactcca cctacagaaa gttttccacc aaagtcacgc tgaatacagt 480
ggggcaccac cacccccccc cgccccatca ggcctccgtt tctggaatcc aagcagaatt 540
gctcacattt cccaactcct ctcctggcct cagaagacag aagagagact gggttattcc 600
tcccatcagc tgcccagaaa atgaaaaagg cccatttcct aaaaacctgg ttcagatcaa 660
atccaacaaa gacaaagaag gcaaggtttt ctacagcatc actggccaag gagctgacac 720
acccctgtt ggtgtcttta ttattgaaag agaaacagga tggctgaagg tgacagagcc 780
tctggataga gaacgcattg ccacatacac tctcttctct cacgctgtgt catccaacgg 840
gaatgcagtt gaggatccaa tggagatttt gatcacggta accgatcaga atgacaacaa 900
gcccgaattc acccaggagg tctttaaggg gtctgtcatg gaaggtgctc ttccaggaac 960
ctctgtgatg gaggtcacag ccacagacgc ggacgatgat gtgaacacct acaatgccgc 1020
catcgcttac accatcctca gccaagatcc tgagctcccc gacaaaaata tgttcaccat 1080
taacaggaac acaggagtca tcagtgtggt caccactggg ctggaccgag agagtttccc 1140
tacgtatacc ctggtggttc aagctgctga ccttcaaggt gaggggttaa gcacaacagc 1200
aacagctgtg atcacagtca ctgacaccaa cgataatcct ccgatcttca atcccaccac 1260
gtacaagggt caggtgcctg agaacgaggc taacgtcgta atcaccacac tgaaagtgac 1320
tgatgctgat gcccccaata ccccagcgtg ggaggctgta tacaccatat tgaatgatga 1380
tggtggacaa tttgtcgtca ccacaaatcc agtgaacaac gatggcattt tgaaaacagc 1440
aaagggcttg gattttgagg ccaagcagca gtacattcta cacgtagcag tgacgaatgt 1500
ggtacctttt gaggtctctc tcaccacctc cacagccacc gtcacccgtgg atgtgctgga 1560
tgtgaatgaa gcccccatct ttgtgcctcc tgaaaagaga gtggaagtgt ccgaggactt 1620
```

tggcgtgggc caggaaatca catcctacac tgcccaggag ccagacacat ttatggaaca 1680
gaaaataaca tatcggattt ggagagacac tgccaactgg ctggagatta atccggacac 1740
tggtgccatt tccactcggg ctgagctgga cagggaggat tttgagcacg tgaagaacag 1800
cacgtacaca gccctaatca tagctacaga caatggttct ccagttgcta ctggaacagg 1860
gacacttctg ctgatcctgt ctgatgtgaa tgacaacgcc cccataccag aacctcgaac 1920
tatattcttc tgtgagagga atccaaagcc tcaggtcata aacatcattg atgcagacct 1980
tcctcccaat acatctccct tcacagcaga actaacacac ggggcgagtg ccaactggac 2040
cattcagtac aacgacccaa cccaagaatc tatcatttta aagccaaaga tggccttaga 2100
ggtgggtgac tacaaaatca atctcaagct catggataac cagaataaag accaagtgac 2160
caccttagag gtcagcgtgt gtgactgtga aggggccgcc ggcgtctgta ggaaggcaca 2220
gcctgtcgaa gcaggattgc aaattcctgc cattctgggg attcttggag gaattcttgc 2280
tttgctaatt ctgattctgc tgctcttgct gtttcttcgg aggagagcgg tggtcaaaga 2340
gcccttactg cccccagagg atgacacccg ggacaacgtt tattactatg atgaagaagg 2400
aggcggagaa gaggaccagg actttgactt gagccagctg cacaggggcc tggacgctcg 2460
gcctgaagtg actcgtaacg acgttgcacc aaccctcatg agtgtccccc ggtatcttcc 2520
ccgcctcgcc aatcccgatg aaattggaaa ttttattgat gaaaatctga aagcggctga 2580
tactgacccc acagccccgc cttatgattc tctgctctgtg tttgactatg aaggaagcgg 2640
ttccgaagct gctagtctga gctccctgaa ctcctcagag tcagacaaag accaggacta 2700
tgactacttg aacgaatggg gcaatcgctt caagaagctg gctgacatgt acggaggcgg 2760
cgaggacgac tagcggactc gagagaggcg ggccccagac ccatgtgctg ggaaatgcag 2820
aaatcacgtt gctggtggtt tttcagctcc cttcccttga gatgagtttc tggggaaaaa 2880
aaagagactg gttagtgatg cagttagtat agctttatac tctctccact ttatagctct 2940
aataagtttg tgttagaaaa gtttcgactt atttcttaaa gctttttttt ttttcccatc 3000
actctttaca tggtggtgat gtccaaaaga tacccaaatt ttaatattcc agaagaacaa 3060
ctttagcatc agaaggttca cccagcacct tgcagatttt cttaaggaat tttgtctcac 3120
ttttaaaaag aaggggagaa gtcagctact ctagttctgt tgttttgtgt atataatttt 3180
ttaaaaaaaa tttgtgtgct tctgctcatt actacactgg tgtgtcccta tgccttttct 3240
ttttttttta agacagggtc tcattctatc ggccaggctg gagtgcagtg gtgcaatcac 3300
agctcactgc agcctgtcc tcccaggctc aagctatcct tgcacctcag cctcccaagt 3360
agctgggacc acaggcatgc accactacgc atgactaatt ttttaaatat tgagacgggg 3420
gtctccctgt gttacccagg ctggtctcaa actcctgggc tcaagtgatc ctcccatctt 3480
ggcctcccag agtattggga ttacagacat gagccactgc acctgcccag ctccccaact 3540
ccctgccatt ttttaagaga cagtttcgct ccatcgccca ggcctgggat gcagtgatgt 3600
gatcatagct cactgtaacc tcaaactctg gggctcaagc agttctccca ccagcctcct 3660
ttttattttt ttgtacagat ggggtcttgc tatgttgccc aagctggtct taaactcctg 3720
gcctcaagca atcctctgc cttggccccc caaagtgctg ggattgtgag catgagctgc 3780
tgtgcccagc ctccatgttt taatatcaac tctcactcct gaattcagtt gctttgccca 3840
agatagggagt tctctgatgc agaaattatt gggctctttt aggggtaagaa gtttgtgtct 3900
ttgtctggcc acatcttgac taggtattgt ctactctgaa gaccttttaat ggcttccctc 3960
tttcatctcc tgagtatgta acttgcaatg ggcagctatc cagtgacttg ttctgagtaa 4020
gtgtgttcat taatgtttat ttagctctga agcaagagtg atatactcca ggacttagaa 4080
tagtgcctaa agtgctgcag ccaaagacag agcggaacta tgaaaagtgg gcttggagat 4140
ggcaggagag cttgtcattg agcctggcaa tttagcaaac tgatgctgag gatgattgag 4200
gtggttctac ctcatctctg aaaattctgg aaggaatgga ggagtctcaa catgtgtttc 4260
tgacacaaga tccgtggttt gtactcaaag cccagaatcc ccaagtgcct gcttttgatg 4320
atgtctacag aaaatgctgg ctgagctgaa cacatttgcc caattccagg tgtgcacaga 4380
aaaccgagaa tattcaaaat tccaaatttt ttcttaggag caagaagaaa atgtggccct 4440
aaaggggctt agttgagggg taggggtag tgaggatctt gatttggatc tcttttatt 4500
taaatgtgaa tttcaacttt tgacaatcaa agaaaagact tttgttgaaa tagctttact 4560
gtttctcaag tgttttggag aaaaaaatca accctgcaat cacttttgg aattgtcttg 4620
atttttggc agttcaagct atatcgaata tagttctgtg tagagaatgt cactgtagtt 4680
ttgagtgtat acatgtgtgg gtgctgataa ttgtgtattt tctttggggg tggaaaagga 4740
aaacaattca agctgagaaa agtattctca aagatgcatt tttataaatt ttattaaaca 4800
attttgttaa accataaaaa aaaaaaaa 4828

<210> 37
<211> 882
<212> PRT

<213> Homo sapiens

<400> 37

```
Met Gly Pro Trp Ser Arg Ser Leu Ser Ala Leu Leu Leu Leu Leu Gln
 1               5                  10                  15
Val Ser Ser Trp Leu Cys Gln Glu Pro Glu Pro Cys His Pro Gly Phe
            20                  25                  30
Asp Ala Glu Ser Tyr Thr Phe Thr Val Pro Arg Arg His Leu Glu Arg
        35                  40                  45
Gly Arg Val Leu Gly Arg Val Asn Phe Glu Asp Cys Thr Gly Arg Gln
    50                  55                  60
Arg Thr Ala Tyr Phe Ser Leu Asp Thr Arg Phe Lys Val Gly Thr Asp
65                  70                  75                  80
Gly Val Ile Thr Val Lys Arg Pro Leu Arg Phe His Asn Pro Gln Ile
                85                  90                  95
His Phe Leu Val Tyr Ala Trp Asp Ser Thr Tyr Arg Lys Phe Ser Thr
            100                 105                 110
Lys Val Thr Leu Asn Thr Val Gly His His His Arg Pro Pro Pro His
        115                 120                 125
Gln Ala Ser Val Ser Gly Ile Gln Ala Glu Leu Leu Thr Phe Pro Asn
    130                 135                 140
Ser Ser Pro Gly Leu Arg Arg Gln Lys Arg Asp Trp Val Ile Pro Pro
145                 150                 155                 160
Ile Ser Cys Pro Glu Asn Glu Lys Gly Pro Phe Pro Lys Asn Leu Val
                165                 170                 175
Gln Ile Lys Ser Asn Lys Asp Lys Glu Gly Lys Val Phe Tyr Ser Ile
            180                 185                 190
Thr Gly Gln Gly Ala Asp Thr Pro Pro Val Gly Val Phe Ile Ile Glu
        195                 200                 205
Arg Glu Thr Gly Trp Leu Lys Val Thr Glu Pro Leu Asp Arg Glu Arg
    210                 215                 220
Ile Ala Thr Tyr Thr Leu Phe Ser His Ala Val Ser Ser Asn Gly Asn
225                 230                 235                 240
Ala Val Glu Asp Pro Met Glu Ile Leu Ile Thr Val Thr Asp Gln Asn
                245                 250                 255
Asp Asn Lys Pro Glu Phe Thr Gln Glu Val Phe Lys Gly Ser Val Met
            260                 265                 270
Glu Gly Ala Leu Pro Gly Thr Ser Val Met Glu Val Thr Ala Thr Asp
        275                 280                 285
Ala Asp Asp Asp Val Asn Thr Tyr Asn Ala Ala Ile Ala Tyr Thr Ile
    290                 295                 300
Leu Ser Gln Asp Pro Glu Leu Pro Asp Lys Asn Met Phe Thr Ile Asn
305                 310                 315                 320
Arg Asn Thr Gly Val Ile Ser Val Val Thr Thr Gly Leu Asp Arg Glu
                325                 330                 335
Ser Phe Pro Thr Tyr Thr Leu Val Val Gln Ala Ala Asp Leu Gln Gly
            340                 345                 350
Glu Gly Leu Ser Thr Thr Ala Thr Ala Val Ile Thr Val Thr Asp Thr
        355                 360                 365
Asn Asp Asn Pro Pro Ile Phe Asn Pro Thr Thr Tyr Lys Gly Gln Val
    370                 375                 380
Pro Glu Asn Glu Ala Asn Val Val Ile Thr Thr Leu Lys Val Thr Asp
385                 390                 395                 400
Ala Asp Ala Pro Asn Thr Pro Ala Trp Glu Ala Val Tyr Thr Ile Leu
                405                 410                 415
Asn Asp Asp Gly Gly Gln Phe Val Val Thr Thr Asn Pro Val Asn Asn
            420                 425                 430
Asp Gly Ile Leu Lys Thr Ala Lys Gly Leu Asp Phe Glu Ala Lys Gln
```

```
       435                 440                 445
Gln Tyr Ile Leu His Val Ala Val Thr Asn Val Val Pro Phe Glu Val
    450                 455                 460
Ser Leu Thr Thr Ser Thr Ala Thr Val Thr Val Asp Val Leu Asp Val
465                 470                 475                 480
Asn Glu Ala Pro Ile Phe Val Pro Pro Glu Lys Arg Val Glu Val Ser
                485                 490                 495
Glu Asp Phe Gly Val Gly Gln Glu Ile Thr Ser Tyr Thr Ala Gln Glu
            500                 505                 510
Pro Asp Thr Phe Met Glu Gln Lys Ile Thr Tyr Arg Ile Trp Arg Asp
        515                 520                 525
Thr Ala Asn Trp Leu Glu Ile Asn Pro Asp Thr Gly Ala Ile Ser Thr
    530                 535                 540
Arg Ala Glu Leu Asp Arg Glu Asp Phe Glu His Val Lys Asn Ser Thr
545                 550                 555                 560
Tyr Thr Ala Leu Ile Ile Ala Thr Asp Asn Gly Ser Pro Val Ala Thr
                565                 570                 575
Gly Thr Gly Thr Leu Leu Leu Ile Leu Ser Asp Val Asn Asp Asn Ala
            580                 585                 590
Pro Ile Pro Glu Pro Arg Thr Ile Phe Phe Cys Glu Arg Asn Pro Lys
        595                 600                 605
Pro Gln Val Ile Asn Ile Ile Asp Ala Asp Leu Pro Asn Thr Ser
    610                 615                 620
Pro Phe Thr Ala Glu Leu Thr His Gly Ala Ser Ala Asn Trp Thr Ile
625                 630                 635                 640
Gln Tyr Asn Asp Pro Thr Gln Glu Ser Ile Ile Leu Lys Pro Lys Met
                645                 650                 655
Ala Leu Glu Val Gly Asp Tyr Lys Ile Asn Leu Lys Leu Met Asp Asn
            660                 665                 670
Gln Asn Lys Asp Gln Val Thr Thr Leu Glu Val Ser Val Cys Asp Cys
        675                 680                 685
Glu Gly Ala Ala Gly Val Cys Arg Lys Ala Gln Pro Val Glu Ala Gly
    690                 695                 700
Leu Gln Ile Pro Ala Ile Leu Gly Ile Leu Gly Gly Ile Leu Ala Leu
705                 710                 715                 720
Leu Ile Leu Ile Leu Leu Leu Leu Leu Phe Leu Arg Arg Arg Ala Val
                725                 730                 735
Val Lys Glu Pro Leu Leu Pro Pro Glu Asp Asp Thr Arg Asp Asn Val
            740                 745                 750
Tyr Tyr Tyr Asp Glu Glu Gly Gly Gly Glu Glu Asp Gln Asp Phe Asp
        755                 760                 765
Leu Ser Gln Leu His Arg Gly Leu Asp Ala Arg Pro Glu Val Thr Arg
    770                 775                 780
Asn Asp Val Ala Pro Thr Leu Met Ser Val Pro Arg Tyr Leu Pro Arg
785                 790                 795                 800
Pro Ala Asn Pro Asp Glu Ile Gly Asn Phe Ile Asp Glu Asn Leu Lys
                805                 810                 815
Ala Ala Asp Thr Asp Pro Thr Ala Pro Pro Tyr Asp Ser Leu Leu Val
            820                 825                 830
Phe Asp Tyr Glu Gly Ser Gly Ser Glu Ala Ala Ser Leu Ser Ser Leu
        835                 840                 845
Asn Ser Ser Glu Ser Asp Lys Asp Gln Asp Tyr Asp Tyr Leu Asn Glu
    850                 855                 860
Trp Gly Asn Arg Phe Lys Lys Leu Ala Asp Met Tyr Gly Gly Gly Glu
865                 870                 875                 880
Asp Asp
```

<210> 38
<211> 4521
<212> DNA
<213> Homo sapiens

<400> 38
acttcattca cttgcaaatc agtgtgtgcc cacaagagcc agctctcccg agcccgtaac 60
cttcgcatcc caagagctgc agtttcagcc gcgacagcaa gaacggcaga gccggccgacc 120
gcggcggccg cggcggcgga ggcaggagca gcctgggccg gtcgcagggt ctccgcgggc 180
gcaggaaggc gagcagagat atcctctgag agccaagcaa agaacattaa ggaaggaagg 240
aggaatgagg ctggatacgg tgcagtgaaa aaggcacttc caagagtggg gcactcacta 300
cgcacagact cgacggtgcc atcagcatga gaacttaccg ctacttcttg ctgctctttt 360
gggtgggcca gccctaccca actctctcaa ctccactatc aaagaggact agtggtttcc 420
cagcaaagaa aagggccctg gagctctctg gaaacagcaa aaatgagctg aaccgttcaa 480
aaaggagctg gatgtggaat cagttctttc tcctggagga atacacagga tccgattatc 540
agtatgtggg caagttacat tcagaccagg atagaggaga tggatcactt aaatatatcc 600
tttcaggaga tggagcagga gatctcttca ttattaatga aaacacaggc gacatacagg 660
ccaccaagag gctggacagg gaagaaaaac ccgtttacat ccttcgagct caagctataa 720
acagaaggac agggagaccc gtggagcccg agtctgaatt catcatcaag 780
tcaatgacaa tgaaccaata ttcaccaagg aggtttacac agccactgtc cctgaaatgt 840
ctgatgtcgg tacatttgtt gtccaagtca ctgcgacgga tgcagatgat ccaacatatg 900
ggaacagtgc taaagttgtc tacagtattc tacagggaca gccctatttt tcagttgaat 960
cagaaacagg tattatcaag acagctttgc tcaacatgga tcgagaaaac agggagcagt 1020
accaagtggt gattcaagcc aaggatatgg gcggccagat gggaggatta tctgggacca 1080
ccaccgtgaa catcacactg actgatgtca acgacaaccc tcacccgattc ccccagagta 1140
cataccagtt taaaactcct gaatcttctc caccggggac accaattggc agaatcaaag 1200
ccagcgacgc tgatgtggga gaaaatgctg aaattgagta cagcatcaca gacggtgagg 1260
ggctggatat gtttgatgtc atcaccgacc aggaaaccca ggaagggatt ataactgtca 1320
aaaagctctt ggactttgaa aagaagaaag tgtataccct taaagtggaa gcctccaatc 1380
cttatgttga gccacgattt ctctacttgg ggccttttcaa agattcagcc acggttagaa 1440
ttgtggtgga ggatgtagat gagccacctg tcttcagcaa actggcctac atcttacaaa 1500
taagagaaga tgctcagata aacaccacaa taggctccgt cacagcccaa gatccagatg 1560
ctgccaggaa tcctgtcaag tactctgtag atcgacacac agatatggac agaatattca 1620
acattgattc tggaaatggt tcgatttttta catcgaaact tcttgaccga gaaacactgc 1680
tatggcacaa cattacagtg atagcaacag agatcaataa tccaaagcaa agtagtcgag 1740
tacctctata tattaaagtt ctagatgtca atgacaacgc cccagaattt gctgagttct 1800
atgaaacttt tgtctgtgaa aaagcaaagg cagatcagtt gattcagacc ctgcatgctg 1860
ttgacaagga tgaccctttat agtggacacc aattttcgtt ttccttggcc cctgaagcag 1920
ccagtggctc aaactttacc attcaagaca acaaagacaa cacggcggga atcttaactc 1980
ggaaaaatgg ctataataga cacgagatga gcacctatct cttgcctgtg gtcatttcag 2040
acaacgacta cccagttcaa agcagcactg ggacagtgac tgtccgggtc tgtgcatgtg 2100
accaccacgg gaacatgcaa tcctgccatg cggaggcgct catccacccc acgggactga 2160
gcacgggggc tctggttgcc atccttctgt gcatcgtgat cctactagtg acagtggtgc 2220
tgtttgcagc tctgaggcgg cagcgaaaaa aagagccttt gatcatttcc aaagaggaca 2280
tcagagataa cattgtcagt tacaacgacg aaggtggtgg agaggaggac acccaggctt 2340
ttgatatcgg cacccctgagg aatcctgaag ccatagagga caacaaatta cgaaggagaca 2400
ttgtgcccga agccctttc ctaccccgac ggactccaac agctcgcgac aacacccgatg 2460
tcagagattt cattaaccaa aggttaaagg aaaatgacac ggaccccact gccccgccat 2520
acgactcctt ggccacttac gcctatgaag gcactggctc cgtggcggat tccctgagct 2580
cgctggagtc agtgaccacg gatgcagatc aagactatga ttacctttagt gactgggcac 2640
ctcgattcaa aaagcttgca gatatgtatg gaggagtgga cagtgacaaa gactcctaat 2700
ctgttgcctt tttcattttc caatacgaca ctgaaatatg tgaagtggct atttctttat 2760
atttatccac tactccgtga aggcttctct gttctacccg ttccaaaagc caatggctgc 2820
agtccgtgtg gatccaatgt tagagacttt tttctagtac acttttatga gcttccaagg 2880
ggcaaatttt tatttttag tgcatccagt taaccaagtc agcccaacag gcaggtgccg 2940
gagggagga cagggaacag tatttccact tgttctcagg gcagcgtgcc cgcttccgct 3000
gtcctggtgt tttactacac tccatgtcag gtcagccaac tgccctaact gtacatttca 3060

caggctaatg ggataaagga ctgtgcttta aagataaaaa tatcatcata gtaaaagaaa 3120
tgagggcata tcggctcaca aagagataaa ctacataggg gtgttttattt gtgtcacaaa 3180
gaatttaaaa taacacttgc ccatgctatt tgttcttcaa gaacttttctc tgccatcaac 3240
tactattcaa aacctcaaat ccacccatat gttaaaattc tcattactct taaggaatag 3300
aagcaaatta aacggtaaca tccaaaagca accacaaacc tagtacgact tcattccttc 3360
cactaactca tagtttgtta tatcctagac tagacatgcg aaagtttgcc tttgtaccat 3420
ataaaggggg agggaaatag ctaataatgt taaccaagga aatatatttt accatacatt 3480
taaagttttg gccaccacat gtatcacggg tcacttgaaa ttctttcagc tatcagtagg 3540
ctaatgtcaa aattgtttaa aaatttcttga aagaattttc ctgagacaaa ttttaacttc 3600
ttgtctatag ttgtcagtat tattctacta tactgtacat gaaagtagca gtgtgaagta 3660
caataattca tattcttcat atccttctta cacgactaag ttgaattagt aaagttagat 3720
taaataaaac ttaaatctca ctctaggagt tcagtggaga ggttagagcc agccacactt 3780
gaacctaata ccctgccctt gacatctgga aacctctaca tatttatata acgtgataca 3840
tttggataaa caacattgag attatgatga aaacctacat attccatgtt tggaagacca 3900
ttggaagagg aaaattggat tcccttaaac aaaagtgttt aagattgtaa ttaaaatgat 3960
agttgatttt caaaagcatt aatttttttt cattgttttt aactttgctt tcatgaccat 4020
cctgccatcc ttgactttga actaatgata aagtaatgat ctcaaactat gacagaaaag 4080
taatgtaaaa tccatccaat ctattatttc tctaattatg caattagcct catagttatt 4140
atccagagga cccaactgaa ctgaactaat ccttctggca gattcaaatc gtttatttca 4200
cacgctgttc taatggcact tatcattaga atcttacctt gtgcagtcat cagaaattcc 4260
agcgtactat aatgaaaaca tccttgtttt gaaaacctaa aagacaggct ctgtatatat 4320
atatacttaa gaatatgctg acttcactta ttagtcttag ggatttattt tcaattaata 4380
ttaatttttct acaaatatatt ttagtgtcat ttccatttgg ggatattgtc atatcagcac 4440
atattttctg tttggaaaca cactgttgtt tagttaagtt ttaaataggt gtattaccca 4500
agaagtaaag atggaaacgt t 4521

<210> 39
<211> 790
<212> PRT
<213> Homo sapiens

<400> 39
Met Arg Thr Tyr Arg Tyr Phe Leu Leu Leu Phe Trp Val Gly Gln Pro
 1   5   10   15
Tyr Pro Thr Leu Ser Thr Pro Leu Ser Lys Arg Thr Ser Gly Phe Pro
            20   25   30
Ala Lys Lys Arg Ala Leu Glu Leu Ser Gly Asn Ser Lys Asn Glu Leu
        35   40   45
Asn Arg Ser Lys Arg Ser Trp Met Trp Asn Gln Phe Phe Leu Leu Glu
    50   55   60
Glu Tyr Thr Gly Ser Asp Tyr Gln Tyr Val Gly Lys Leu His Ser Asp
65   70   75   80
Gln Asp Arg Gly Asp Gly Ser Leu Lys Tyr Ile Leu Ser Gly Asp Gly
                85   90   95
Ala Gly Asp Leu Phe Ile Ile Asn Glu Asn Thr Gly Asp Ile Gln Ala
            100   105   110
Thr Lys Arg Leu Asp Arg Glu Glu Lys Pro Val Tyr Ile Leu Arg Ala
        115   120   125
Gln Ala Ile Asn Arg Arg Thr Gly Arg Pro Val Glu Pro Glu Ser Glu
    130   135   140
Phe Ile Ile Lys Ile His Asp Ile Asn Asp Asn Glu Pro Ile Phe Thr
145   150   155   160
Lys Glu Val Tyr Thr Ala Thr Val Pro Glu Met Ser Asp Val Gly Thr
                165   170   175
Phe Val Val Gln Val Thr Ala Thr Asp Ala Asp Asp Pro Thr Tyr Gly
            180   185   190
Asn Ser Ala Lys Val Val Tyr Ser Ile Leu Gln Gly Gln Pro Tyr Phe
        195   200   205

```
Ser Val Glu Ser Glu Thr Gly Ile Ile Lys Thr Ala Leu Leu Asn Met
    210                 215                 220
Asp Arg Glu Asn Arg Glu Gln Tyr Gln Val Val Ile Gln Ala Lys Asp
225                 230                 235                 240
Met Gly Gly Gln Met Gly Gly Leu Ser Gly Thr Thr Thr Val Asn Ile
                245                 250                 255
Thr Leu Thr Asp Val Asn Asp Asn Pro Pro Arg Phe Pro Gln Ser Thr
            260                 265                 270
Tyr Gln Phe Lys Thr Pro Glu Ser Ser Pro Pro Gly Thr Pro Ile Gly
        275                 280                 285
Arg Ile Lys Ala Ser Asp Ala Asp Val Gly Glu Asn Ala Glu Ile Glu
    290                 295                 300
Tyr Ser Ile Thr Asp Gly Glu Gly Leu Asp Met Phe Asp Val Ile Thr
305                 310                 315                 320
Asp Gln Glu Thr Gln Glu Gly Ile Ile Thr Val Lys Lys Leu Leu Asp
                325                 330                 335
Phe Glu Lys Lys Lys Val Tyr Thr Leu Lys Val Glu Ala Ser Asn Pro
            340                 345                 350
Tyr Val Glu Pro Arg Phe Leu Tyr Leu Gly Pro Phe Lys Asp Ser Ala
        355                 360                 365
Thr Val Arg Ile Val Val Glu Asp Val Asp Glu Pro Pro Val Phe Ser
    370                 375                 380
Lys Leu Ala Tyr Ile Leu Gln Ile Arg Glu Asp Ala Gln Ile Asn Thr
385                 390                 395                 400
Thr Ile Gly Ser Val Thr Ala Gln Asp Pro Asp Ala Ala Arg Asn Pro
                405                 410                 415
Val Lys Tyr Ser Val Asp Arg His Thr Asp Met Asp Arg Ile Phe Asn
            420                 425                 430
Ile Asp Ser Gly Asn Gly Ser Ile Phe Thr Ser Lys Leu Leu Asp Arg
        435                 440                 445
Glu Thr Leu Leu Trp His Asn Ile Thr Val Ile Ala Thr Glu Ile Asn
    450                 455                 460
Asn Pro Lys Gln Ser Ser Arg Val Pro Leu Tyr Ile Lys Val Leu Asp
465                 470                 475                 480
Val Asn Asp Asn Ala Pro Glu Phe Ala Glu Phe Tyr Glu Thr Phe Val
                485                 490                 495
Cys Glu Lys Ala Lys Ala Asp Gln Leu Ile Gln Thr Leu His Ala Val
            500                 505                 510
Asp Lys Asp Asp Pro Tyr Ser Gly His Gln Phe Ser Phe Ser Leu Ala
        515                 520                 525
Pro Glu Ala Ala Ser Gly Ser Asn Phe Thr Ile Gln Asp Asn Lys Asp
    530                 535                 540
Asn Thr Ala Gly Ile Leu Thr Arg Lys Asn Gly Tyr Asn Arg His Glu
545                 550                 555                 560
Met Ser Thr Tyr Leu Leu Pro Val Val Ile Ser Asp Asn Asp Tyr Pro
                565                 570                 575
Val Gln Ser Ser Thr Gly Thr Val Thr Val Arg Val Cys Ala Cys Asp
            580                 585                 590
His His Gly Asn Met Gln Ser Cys His Ala Glu Ala Leu Ile His Pro
        595                 600                 605
Thr Gly Leu Ser Thr Gly Ala Leu Val Ala Ile Leu Leu Cys Ile Val
    610                 615                 620
Ile Leu Leu Val Thr Val Val Leu Phe Ala Ala Leu Arg Arg Gln Arg
625                 630                 635                 640
Lys Lys Glu Pro Leu Ile Ile Ser Lys Glu Asp Ile Arg Asp Asn Ile
                645                 650                 655
Val Ser Tyr Asn Asp Glu Gly Gly Gly Glu Glu Asp Thr Gln Ala Phe
            660                 665                 670
```

```
Asp Ile Gly Thr Leu Arg Asn Pro Glu Ala Ile Glu Asp Asn Lys Leu
        675                 680                 685
Arg Arg Asp Ile Val Pro Glu Ala Leu Phe Leu Pro Arg Arg Thr Pro
    690                 695                 700
Thr Ala Arg Asp Asn Thr Asp Val Arg Asp Phe Ile Asn Gln Arg Leu
705                 710                 715                 720
Lys Glu Asn Asp Thr Asp Pro Thr Ala Pro Pro Tyr Asp Ser Leu Ala
                725                 730                 735
Thr Tyr Ala Tyr Glu Gly Thr Gly Ser Val Ala Asp Ser Leu Ser Ser
            740                 745                 750
Leu Glu Ser Val Thr Thr Asp Ala Asp Gln Asp Tyr Asp Tyr Leu Ser
        755                 760                 765
Asp Trp Gly Pro Arg Phe Lys Lys Leu Ala Asp Met Tyr Gly Gly Val
    770                 775                 780
Asp Ser Asp Lys Asp Ser
785                 790
```

<210> 40
<211> 987
<212> DNA
<213> Homo sapiens

<400> 40
cggagagggg gagaacagac aacgggcggc ggggagcagc atggagccgg cggcggggag 60
cagcatggag ccttcggctg actggctggc cacggccgcg gcccggggtc gggtagagga 120
ggtgcggggcg ctgctggagg cgggggcgct gcccaacgca ccgaatagtt acggtcggag 180
gccgatccag gtcatgatga tgggcagcgc ccgagtggcg gagctgctgc tgctccacgg 240
cgcggagccc aactgcgccg accccgccac tctcacccga cccgtgcacg acgctgcccg 300
ggagggcttc ctggacacgc tggtggtgct gcaccgggcc ggggcgcggc tggacgtgcg 360
cgatgcctgg ggccgtctgc ccgtggacct ggctgaggag ctgggccatc gcgatgtcgc 420
acggtacctg cgcgcggctg cgggggcgac cagaggcagt aaccatgccc gcatagatgc 480
cgcggaaggt ccctcagaca tccccgattg aaagaaccag agaggctctg agaaacctcg 540
ggaaacttag atcatcagtc accgaaggtc ctacagggcc acaactgccc ccgccacaac 600
ccaccccgct ttcgtagttt tcatttagaa aatagagctt ttaaaaatgt cctgcctttt 660
aacgtagata taagccttcc cccactaccg taaatgtcca tttatatcat tttttatata 720
ttcttataaa aatgtaaaaa agaaaaacac cgcttctgcc ttttcactgt gttggagttt 780
tctggagtga gcactcacgc cctaagcgca cattcatgtg ggcatttctt gcgagcctcg 840
cagcctccgg aagctgtcga cttcatgaca agcattttgt gaactaggga agctcagggg 900
ggttactggc ttctcttgag tcacactgct agcaaatggc agaaccaaag ctcaaataaa 960
aataaaataa ttttcattca ttcactc 987

<210> 41
<211> 156
<212> PRT
<213> Homo sapiens

<400> 41
```
Met Glu Pro Ala Ala Gly Ser Ser Met Glu Pro Ser Ala Asp Trp Leu
 1               5                  10                  15
Ala Thr Ala Ala Ala Arg Gly Arg Val Glu Glu Val Arg Ala Leu Leu
            20                  25                  30
Glu Ala Gly Ala Leu Pro Asn Ala Pro Asn Ser Tyr Gly Arg Arg Pro
        35                  40                  45
Ile Gln Val Met Met Met Gly Ser Ala Arg Val Ala Glu Leu Leu Leu
    50                  55                  60
Leu His Gly Ala Glu Pro Asn Cys Ala Asp Pro Ala Thr Leu Thr Arg
65                  70                  75                  80
```

```
Pro Val His Asp Ala Ala Arg Glu Gly Phe Leu Asp Thr Leu Val Val
                85                  90                  95
Leu His Arg Ala Gly Ala Arg Leu Asp Val Arg Asp Ala Trp Gly Arg
            100                 105                 110
Leu Pro Val Asp Leu Ala Glu Glu Leu Gly His Arg Asp Val Ala Arg
        115                 120                 125
Tyr Leu Arg Ala Ala Ala Gly Gly Thr Arg Gly Ser Asn His Ala Arg
    130                 135                 140
Ile Asp Ala Ala Glu Gly Pro Ser Asp Ile Pro Asp
145                 150                 155


<210> 42
<211> 5142
<212> DNA
<213> Homo sapiens

<400> 42
gaattcggcc gagaggacga gggggagggc cagagctgcg cgtgctgctt tgcccgagcc 60
cgagcccgag cccgagccag agcccgagcc cgagcccgag cgaacgcca agcctgggag 120
cgcggagccc ggctagggac tcctcctatt tatggagcag gcacccaaca tggctgagcc 180
ccggggcccc gtagaccatg gagtccagat tcgcttcatc acagagccag tgagtggtgc 240
agagatgggc actctacgtc gaggtggacg acgcccagct aaggatgcaa gagccagtac 300
ctacgggggtt gctgtgcgtg tgcagggaat cgctgggcag ccctttgtgg tgctcaacag 360
tggggagaaa ggcggtgact cctttggggt ccaaatcaag ggggccaatg accaaggggc 420
ctcaggagct ctgagctcag atttggaact ccctgagaac ccctactctc aggtcaaggg 480
atttcctgcc ccctcgcaga gcagcacatc tgatgaggag cctgggggct actggaatgg 540
aaagctactc cgttcccact cccaggcctc actggcaggc cctgggccag tggatcctag 600
taacagaagc aacagcatgc tggagctagc cccgaaagtg gcttccccag gtagcaccat 660
tgacactgct cccctgtctt cagtggactc actcatcaac aagtttgaca gtcaacttgg 720
aggccaggcc cggggtggga ctggccgccg aacacggatg ctaccccctg aacagcgcaa 780
acggagcaag agcctggaca gccgcctccc acgggacacc tttgaggaac gggagcgcca 840
gtccaccaac cactggacct ctagcacaaa atatgacaac catgtgggca cttcgaagca 900
gccagcccag agccagaacc tgagtcctct cagtggcttt gtcgagctca gtcagactca 960
ggactgggtc cttcagagtt ttgaggagcc gcggaggagt gcacaggacc ccaccatgct 1020
gcagttcaaa tcaactccag acctccttcg agaccagcag gaggcagccc caccaggcag 1080
tgtggaccat atgaaggcca ccatctatgg catcctgagg gagggaagct cagaaagtga 1140
aacctctgtg aggaggaagg ttagtttggt gctggagaag atgcagcctc tagtgatggt 1200
ttcttctggt tctactaagg ccgtggcagg gcaggggtgag cttacccgaa aagtggagga 1260
gctacagcga aagctggatg aagaggtgaa gaagcggcag aagctagagc catcccaagt 1320
tgggctggag cggcagctgg aggagaaaac agaagaqtgc agccgactgc aggagctgct 1380
ggagaggagg aaggggagg cccagcagag caacaaggag ctccagcaca tgaagcgcct 1440
cttggaccag ggtgaagatt tacgacatgg gctggagacc caggtgatgg agctgcagaa 1500
caagctgaaa catgtccagg gtcctgagcc tgctaaggag gtgttactga agcaqcctc 1560
agagacccgg gaacttctgg aagaggtctt ggaggggaaa cagcgagtag aggagcagct 1620
gaggctgcgg gagcgggagt tgacagccct gaaggggcc ctgaaagagg agctagcctc 1680
ccgtgaccag gaggtggaac atgtccggca gcagtaccag cgagacacag agcagctccg 1740
caggagcatg caagatgcaa cccaggacca tgcaglgctg gaggccgaga ggcagaagat 1800
gtcagccctt gtgcgagggc tgcagaggga gctggaggag acttcagagg agacagggca 1860
ttggcagagt atgttccaga agaacaagga ggatcttaga gccaccaagc aggaactcct 1920
gcagctgcga atggagaagg aggagatgga agaggagctt ggagagaaga tagaggtctt 1980
gcagagggaa ttagagcagg cccgagctag tgctggagat actcgccagg ttgaggtgct 2040
caagaaggag ctgctccgga cacaggagga gcttaaggaa ctgcaggcag aacggcagag 2100
ccaggaggtg gctgggcgac accgggaccg ggagttggag aagcagctgg cggtcctgag 2160
ggtcgaggct gatcgaggtc gggagctgga agaacagaac ctccagctac aaaagaccct 2220
ccagcaactg cgacaggact gtgaagaggc ttccaaggct aagatggtgg ccgaggcaga 2280
ggcaacagtg ctggggcagc ggcgggccgc agtggagacg acgcttcggg agaccccagga 2340
ggaaaatgac gaattccgcc ggcgcatcct gggtttggag cagcagctga aggagactcg 2400
```

```
aggtctggtg gatggtgggg aagcggtgga ggcacgacta cgggacaagc tgcagcggct 2460
ggaggcagag aaacagcagc tggaggaggc cctgaatgcg tcccaggaag aggaggggag 2520
tctggcagca gccaagcggg cactggaggc acgcctagag gaggctcagc gggggctggc 2580
ccgcctgggg caggagcagc agacactgaa ccgggccctg gaggaggaag ggaagcagcg 2640
ggaggtgctc cggcgaggca aggctgagct ggaggagcag aagcgtttgc tggacaggac 2700
tgtggaccga ctgaacaagg agttggagaa gatcggggag gactctaagc aagccctgca 2760
gcagctccag gcccagctgg aggattataa ggaaaagccc cggcgggagg tggcagatgc 2820
ccagcgccag gccaaggatt gggccagtga ggctgagaag acctctggag gactgagccg 2880
acttcaggat gagatccaga ggctgcggca ggccctgcag gcatcccagg ctgagcggga 2940
cacagcccgg ctggacaaag agctactggc ccagcgactg cagggggctgg agcaagaggc 3000
agagaacaag aagcgttccc aggacgacag ggccccggcag ctgaagggtc tcgaggaaaa 3060
agtctcacgg ctggaaacag agttagatga ggagaagaac accgtggagc tgctaacaga 3120
tcgggtgaat cgtggccggg accaggtgga tcagctgagg acagagctca tgcaggaaag 3180
gtctgctcgg caggacctgg agtgtgacaa aatctccttg gagagacaga acaaggacct 3240
gaagacccgg ttggccagct cagaaggctt ccagaaqcct agtgccagcc tctctcagct 3300
tgagtcccag aatcagttgt tgcaggagcg gctacaggct gaagagaggg agaagacagt 3360
tctgcagtct accaatcgaa aactggagcg gaaagttaaa gaactatcca tccagattga 3420
agacgagcgg cagcatgtca atgaccagaa agaccagcta agcctgaggg tgaaggcttt 3480
gaagcgtcag gtggatgaag cagaagagga aattgagcga ctggacggcc tgaggaagaa 3540
ggcccagcgt gaggtggagg agcagcatga ggtcaatgaa cagctccagg cccggatcaa 3600
gtctctggag aaggactcct ggcgcaaagc ttcccgctca gctgctgagt cagctctcaa 3660
aaacgaaggg ctgagctcag atgaggaatt cgacagtgtc tacgatccct cgtccattgc 3720
atcactgctt acggagagca acctacagac cagctcctgt tagctcgtgg tcctcaagga 3780
ctcagaaacc aggctcgagg cctatcccag caagtgctgc tctgctctgc ccaccctggg 3840
ttctgcattc ctatgggtga cccaattatt cagacctaag acaggtaggg gtcagagtga 3900
tggtgataaa aaaaaaaaaa tcatcagcaa taagctgata gatggacttt ccactgtagg 3960
agtggacgtt tcaagccaac tgagcctttt cctcaagtgc cgacacctcc ctcatctctc 4020
ttatagtgga aggatggtca gcattaggct gatggggact gagaaggata ggaagggata 4080
gaaattgcca tgtgtataaa gctttattct ttagccctta accctaaggc tcaggtaaat 4140
accctatgtt attgtgctcc ctggattcct gcaactcatt ttccttccac tctggagcag 4200
ggtgagggga atgttatggg taacagacat gcaggcatgg ctctacccat ttctttgcac 4260
aagtatgggg cccatgtggt agtccccata cccctccagt tcctatattt ttgtcttctt 4320
cctttccccc ctttgccatt cctaccttgc atttttcctg tcagtgcctt agccaaggca 4380
aggagataag gatgctcttc ttgcttttta tatctgcaca ttcatacctc tccaaagacc 4440
agcttttccc cagccagggc cctcagcctt ccctgctgcc ccagtgattg attgagagag 4500
ctgttggggt ttctctgcca atgaccccctg ggagagggac tttggtaggg tcatgataaa 4560
gtggcggggg tcttggtcct gctcagggtt ttcatccttc ctcctctccc tcctctgtga 4620
ctgtggatat ggttataagg tggttgcacc tgggagccct gacaactggc tgcacaaatt 4680
ccaaaagtaa aggtgtcagt ccctgtggcc ttcctgggg cttctctgac cacatgtgcc 4740
caacttcaat aagagaacca agggaccttc attttctgag gtgcttggct ctgattcagg 4800
gctttgcaag gggttagaag ctgactgtaa aaatgggaag aggcaacgga agacatttat 4860
ttctcctttg gattttgggg agaaccaagc cctggtaggg aagaggtaag ggggatgatt 4920
cacctccata tttcctaagc aggttgtata gggagccggt ggcaggagga aggctgtttt 4980
cacaaatgac ttgtaatgtc gtgattaaaa aaattcctat attcttctgc aaatcaaacg 5040
ttctttccca atccaatcca gccttggttt tattttaaat taaatattaa aattacacat 5100
ttatattgaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa                    5142

<210> 43
<211> 1203
<212> PRT
<213> Homo sapiens

<400> 43
Met Glu Gln Ala Pro Asn Met Ala Glu Pro Arg Gly Pro Val Asp His
 1               5                  10                  15
Gly Val Gln Ile Arg Phe Ile Thr Glu Pro Val Ser Gly Ala Glu Met
            20                  25                  30
Gly Thr Leu Arg Arg Gly Gly Arg Arg Pro Ala Lys Asp Ala Arg Ala
```

```
        35                  40                  45
Ser Thr Tyr Gly Val Ala Val Arg Val Gln Gly Ile Ala Gly Gln Pro
    50                  55                  60
Phe Val Val Leu Asn Ser Gly Glu Lys Gly Gly Asp Ser Phe Gly Val
65                  70                  75                  80
Gln Ile Lys Gly Ala Asn Asp Gln Gly Ala Ser Gly Ala Leu Ser Ser
                85                  90                  95
Asp Leu Glu Leu Pro Glu Asn Pro Tyr Ser Gln Val Lys Gly Phe Pro
            100                 105                 110
Ala Pro Ser Gln Ser Ser Thr Ser Asp Glu Glu Pro Gly Ala Tyr Trp
        115                 120                 125
Asn Gly Lys Leu Leu Arg Ser His Ser Gln Ala Ser Leu Ala Gly Pro
    130                 135                 140
Gly Pro Val Asp Pro Ser Asn Arg Ser Asn Ser Met Leu Glu Leu Ala
145                 150                 155                 160
Pro Lys Val Ala Ser Pro Gly Ser Thr Ile Asp Thr Ala Pro Leu Ser
                165                 170                 175
Ser Val Asp Ser Leu Ile Asn Lys Phe Asp Ser Gln Leu Gly Gly Gln
            180                 185                 190
Ala Arg Gly Arg Thr Gly Arg Arg Thr Arg Met Leu Pro Pro Glu Gln
        195                 200                 205
Arg Lys Arg Ser Lys Ser Leu Asp Ser Arg Leu Pro Arg Asp Thr Phe
    210                 215                 220
Glu Glu Arg Glu Arg Gln Ser Thr Asn His Trp Thr Ser Ser Thr Lys
225                 230                 235                 240
Tyr Asp Asn His Val Gly Thr Ser Lys Gln Pro Ala Gln Ser Gln Asn
                245                 250                 255
Leu Ser Pro Leu Ser Gly Phe Ser Arg Ser Arg Gln Thr Gln Asp Trp
            260                 265                 270
Val Leu Gln Ser Phe Glu Glu Pro Arg Arg Ser Ala Gln Asp Pro Thr
        275                 280                 285
Met Leu Gln Phe Lys Ser Thr Pro Asp Leu Leu Arg Asp Gln Gln Glu
    290                 295                 300
Ala Ala Pro Pro Gly Ser Val Asp His Met Lys Ala Thr Ile Tyr Gly
305                 310                 315                 320
Ile Leu Arg Glu Gly Ser Ser Glu Ser Glu Thr Ser Val Arg Arg Lys
                325                 330                 335
Val Ser Leu Val Leu Glu Lys Met Gln Pro Leu Val Met Val Ser Ser
            340                 345                 350
Gly Ser Thr Lys Ala Val Ala Gly Gln Gly Glu Leu Thr Arg Lys Val
        355                 360                 365
Glu Glu Leu Gln Arg Lys Leu Asp Glu Glu Val Lys Lys Arg Gln Lys
    370                 375                 380
Leu Glu Pro Ser Gln Val Gly Leu Glu Arg Gln Leu Glu Glu Lys Thr
385                 390                 395                 400
Glu Glu Cys Ser Arg Leu Gln Glu Leu Leu Glu Arg Arg Lys Gly Glu
                405                 410                 415
Ala Gln Gln Ser Asn Lys Glu Leu Gln Asn Met Lys Arg Leu Leu Asp
            420                 425                 430
Gln Gly Glu Asp Leu Arg His Gly Leu Glu Thr Gln Val Met Glu Leu
        435                 440                 445
Gln Asn Lys Leu Lys His Val Gln Gly Pro Glu Pro Ala Lys Glu Val
    450                 455                 460
Leu Leu Lys Asp Leu Leu Glu Thr Arg Glu Leu Leu Glu Glu Val Leu
465                 470                 475                 480
Glu Gly Lys Gln Arg Val Glu Glu Gln Leu Arg Leu Arg Glu Arg Glu
                485                 490                 495
Leu Thr Ala Leu Lys Gly Ala Leu Lys Glu Glu Val Ala Ser Arg Asp
```

```
            500                 505                 510
Gln Glu Val Glu His Val Arg Gln Gln Tyr Gln Arg Asp Thr Glu Gln
        515                 520                 525
Leu Arg Arg Ser Met Gln Asp Ala Thr Gln Asp His Ala Val Leu Glu
    530                 535                 540
Ala Glu Arg Gln Lys Met Ser Ala Leu Val Arg Gly Leu Gln Arg Glu
545                 550                 555                 560
Leu Glu Glu Thr Ser Glu Glu Thr Gly His Trp Gln Ser Met Phe Gln
                565                 570                 575
Lys Asn Lys Glu Asp Leu Arg Ala Thr Lys Gln Glu Leu Leu Gln Leu
            580                 585                 590
Arg Met Glu Lys Glu Glu Met Glu Glu Glu Leu Gly Glu Lys Ile Glu
        595                 600                 605
Val Leu Gln Arg Glu Leu Glu Gln Ala Arg Ala Ser Ala Gly Asp Thr
    610                 615                 620
Arg Gln Val Glu Val Leu Lys Lys Glu Leu Leu Arg Thr Gln Glu Glu
625                 630                 635                 640
Leu Lys Glu Leu Gln Ala Glu Arg Gln Ser Gln Glu Val Ala Gly Arg
                645                 650                 655
His Arg Asp Arg Glu Leu Glu Lys Gln Leu Ala Val Leu Arg Val Glu
            660                 665                 670
Ala Asp Arg Gly Arg Glu Leu Glu Glu Gln Asn Leu Gln Leu Gln Lys
        675                 680                 685
Thr Leu Gln Gln Leu Arg Gln Asp Cys Glu Glu Ala Ser Lys Ala Lys
    690                 695                 700
Met Val Ala Glu Ala Glu Ala Thr Val Leu Gly Gln Arg Arg Ala Ala
705                 710                 715                 720
Val Glu Thr Thr Leu Arg Glu Thr Gln Glu Glu Asn Asp Glu Phe Arg
                725                 730                 735
Arg Arg Ile Leu Gly Leu Glu Gln Gln Leu Lys Glu Thr Arg Gly Leu
            740                 745                 750
Val Asp Gly Gly Glu Ala Val Glu Ala Arg Leu Arg Asp Lys Leu Gln
        755                 760                 765
Arg Leu Glu Ala Glu Lys Gln Gln Leu Glu Glu Ala Leu Asn Ala Ser
    770                 775                 780
Gln Glu Glu Glu Gly Ser Leu Ala Ala Ala Lys Arg Ala Leu Glu Ala
785                 790                 795                 800
Arg Leu Glu Glu Ala Gln Arg Gly Leu Ala Arg Leu Gly Gln Glu Gln
                805                 810                 815
Gln Thr Leu Asn Arg Ala Leu Glu Glu Glu Gly Lys Gln Arg Glu Val
            820                 825                 830
Leu Arg Arg Gly Lys Ala Glu Leu Glu Glu Gln Lys Arg Leu Leu Asp
        835                 840                 845
Arg Thr Val Asp Arg Leu Asn Lys Glu Leu Glu Lys Ile Gly Glu Asp
    850                 855                 860
Ser Lys Gln Ala Leu Gln Gln Leu Gln Ala Gln Leu Glu Asp Tyr Lys
865                 870                 875                 880
Glu Lys Ala Arg Arg Glu Val Ala Asp Ala Gln Arg Gln Ala Lys Asp
                885                 890                 895
Trp Ala Ser Glu Ala Glu Lys Thr Ser Gly Gly Leu Ser Arg Leu Gln
            900                 905                 910
Asp Glu Ile Gln Arg Leu Arg Gln Ala Leu Gln Ala Ser Gln Ala Glu
        915                 920                 925
Arg Asp Thr Ala Arg Leu Asp Lys Glu Leu Leu Ala Gln Arg Leu Gln
    930                 935                 940
Gly Leu Glu Gln Glu Ala Glu Asn Lys Lys Arg Ser Gln Asp Asp Arg
945                 950                 955                 960
Ala Arg Gln Leu Lys Gly Leu Glu Glu Lys Val Ser Arg Leu Glu Thr
```

```
                965                 970                 975
Glu Leu Asp Glu Glu Lys Asn Thr Val Glu Leu Leu Thr Asp Arg Val
            980                 985                 990
Asn Arg Gly Arg Asp Gln Val Asp Gln Leu Arg Thr Glu Leu Met Gln
        995                 1000                1005
Glu Arg Ser Ala Arg Gln Asp Leu Glu Cys Asp Lys Ile Ser Leu Glu
    1010                1015                1020
Arg Gln Asn Lys Asp Leu Lys Thr Arg Leu Ala Ser Ser Glu Gly Phe
1025                1030                1035                1040
Gln Lys Pro Ser Ala Ser Leu Ser Gln Leu Glu Ser Gln Asn Gln Leu
                1045                1050                1055
Leu Gln Glu Arg Leu Gln Ala Glu Glu Arg Glu Lys Thr Val Leu Gln
            1060                1065                1070
Ser Thr Asn Arg Lys Leu Glu Arg Lys Val Lys Glu Leu Ser Ile Gln
        1075                1080                1085
Ile Glu Asp Glu Arg Gln His Val Asn Asp Gln Lys Asp Gln Leu Ser
    1090                1095                1100
Leu Arg Val Lys Ala Leu Lys Arg Gln Val Asp Glu Ala Glu Glu Glu
1105                1110                1115                1120
Ile Glu Arg Leu Asp Gly Leu Arg Lys Lys Ala Gln Arg Glu Val Glu
                1125                1130                1135
Glu Gln His Glu Val Asn Glu Gln Leu Gln Ala Arg Ile Lys Ser Leu
            1140                1145                1150
Glu Lys Asp Ser Trp Arg Lys Ala Ser Arg Ser Ala Ala Glu Ser Ala
        1155                1160                1165
Leu Lys Asn Glu Gly Leu Ser Ser Asp Glu Glu Phe Asp Ser Val Tyr
    1170                1175                1180
Asp Pro Ser Ser Ile Ala Ser Leu Leu Thr Glu Ser Asn Leu Gln Thr
1185                1190                1195                1200
Ser Ser Cys
```

<210> 44
<211> 1925
<212> DNA
<213> Homo sapiens

<400> 44

```
agtggagtgg gacaggtata taaaggaagt acagggcctg gggaagaggc cctgtctagg 60
tagctggcac caggagccgt gggcaaggga agaggccaca ccctgccctg ctctgctgca 120
gccagaatgg gtgtgaaggc gtctcaaaca ggctttgtgg tcctggtgct gctccagtgc 180
tgctctgcat acaaactggt ctgctactac accagctggt cccagtaccg ggaaggcgat 240
gggagctgct tcccagatgc ccttgaccgc ttcctctgta cccacatcat ctacagcttt 300
gccaatataa gcaacgatca catcgacacc tgggagtgga atgatgtgac gctctacggc 360
atgctcaaca cactcaagaa caggaacccc aacctgaaga ctctcttgtc tgtcggagga 420
tggaactttg ggtctcaaag attttccaag atagcctcca acacccagag tcgccggact 480
ttcatcaagt cagtaccgcc attcctgcgc acccatggct ttgatgggct ggaccttgcc 540
tggctctacc ctggacggag agacaaacag cattttacca ccctaatcaa ggaaatgaag 600
gccgaattta taaaggaagc ccagccaggg aaaaagcagc tcctgatcag cgcagcactg 660
tctgcgggga aggtcaccat tgacagcagc tatgacattg ccaagatatc ccaacacctg 720
gatttcatta gcatcatgac ctacgatttt catggagcct ggcgtgggac cacaggccat 780
cacagtcccc tgttccgagg tcaggaggat gcaagttcctg acagattcag caacactgac 840
tatgctgtgg ggtacatgtt gaggctgggg gctcctgcca gtaagctggt gatgggcatc 900
cccaccttcg ggaggagctt cactctggct tcttctgaga ctggtgttgg agccccaatc 960
tcaggaccgg gaattccagg ccggttcacc aaggaggcag ggacccttgc ctactatgag 1020
atctgtgact tcctccgcgg agccacagtc catagaaccc tcggccagca ggtcccctat 1080
gccaccaagg gcaaccagtg ggtaggatac gacgaccagg aaagcgtcaa aagcaaggtg 1140
```

cagtacctga aggataggca gctggcaggc gccatggtat gggccctgga cctggatgac 1200
ttccagggct ccttctgcgg ccaggatctg cgcttccctc tcaccaatgc catcaaggat 1260
gcactcgctg caacgtagcc ctctgttctg cacacagcac gggggccaag gatgccccgt 1320
cccctatgg ctccagctgg ccgggagcct gatcacctgc cctgctgagt cccaggctga 1380
gcctcagtct ccctcccttg gggcctatgc agaggtccac aacacacaga tttgagctca 1440
gccctggtgg gcagagaggt agggatgggg ctgtggggat agtgaggcat cgcaatgtaa 1500
gactogggat tagtacacac ttgttgatga ttaatggaaa tgtttacaga tccccaagcc 1560
tggcaaggga atttcttcaa ctccctgccc cctagccctc cttatcaaag gacaccattt 1620
tggcaagctc tatcaccaag gagccaaaca tcctacaaga cacagtgacc atactaatta 1680
taccccctgc aaagccagct tgaaaccttc acttaggaac gtaatcgtgt cccctatcct 1740
acttcccctt cctaattcca cagctgctca ataaagtaca agagtttaac agtgtgttgg 1800
cgctttgctt tggtctatct ttgagcgccc actagaccca ctggactcac ctcccccatc 1860
tcttctgggt tccttcctct gagccttggg acccctgagc ttgcagagat gaaggccgcc 1920
atgtt 1925

<210> 45
<211> 383
<212> PRT
<213> Homo sapiens

<400> 45
Met Gly Val Lys Ala Ser Gln Thr Gly Phe Val Val Leu Val Leu Leu
1 5 10 15
Gln Cys Cys Ser Ala Tyr Lys Leu Val Cys Tyr Tyr Thr Ser Trp Ser
20 25 30
Gln Tyr Arg Glu Gly Asp Gly Ser Cys Phe Pro Asp Ala Leu Asp Arg
35 40 45
Phe Leu Cys Thr His Ile Ile Tyr Ser Phe Ala Asn Ile Ser Asn Asp
50 55 60
His Ile Asp Thr Trp Glu Trp Asn Asp Val Thr Leu Tyr Gly Met Leu
65 70 75 80
Asn Thr Leu Lys Asn Arg Asn Pro Asn Leu Lys Thr Leu Leu Ser Val
85 90 95
Gly Gly Trp Asn Phe Gly Ser Gln Arg Phe Ser Lys Ile Ala Ser Asn
100 105 110
Thr Gln Ser Arg Arg Thr Phe Ile Lys Ser Val Pro Pro Phe Leu Arg
115 120 125
Thr His Gly Phe Asp Gly Leu Asp Leu Ala Trp Leu Tyr Pro Gly Arg
130 135 140
Arg Asp Lys Gln His Phe Thr Thr Leu Ile Lys Glu Met Lys Ala Glu
145 150 155 160
Phe Ile Lys Glu Ala Gln Pro Gly Lys Lys Gln Leu Leu Leu Ser Ala
165 170 175
Ala Leu Ser Ala Gly Lys Val Thr Ile Asp Ser Ser Tyr Asp Ile Ala
180 185 190
Lys Ile Ser Gln His Leu Asp Phe Ile Ser Ile Met Thr Tyr Asp Phe
195 200 205
His Gly Ala Trp Arg Gly Thr Thr Gly His His Scr Pro Leu Phe Arg
210 215 220
Gly Gln Glu Asp Ala Ser Pro Asp Arg Phe Ser Asn Thr Asp Tyr Ala
225 230 235 240
Val Gly Tyr Met Leu Arg Leu Gly Ala Pro Ala Ser Lys Leu Val Met
245 250 255
Gly Ile Pro Thr Phe Gly Arg Ser Phe Thr Leu Ala Ser Ser Glu Thr
260 265 270
Gly Val Gly Ala Pro Ile Ser Gly Pro Gly Ile Pro Gly Arg Phe Thr
275 280 285
Lys Glu Ala Gly Thr Leu Ala Tyr Tyr Glu Ile Cys Asp Phe Leu Arg

```
    290                 295                 300
Gly Ala Thr Val His Arg Thr Leu Gly Gln Gln Val Pro Tyr Ala Thr
305                 310                 315                 320
Lys Gly Asn Gln Trp Val Gly Tyr Asp Asp Gln Glu Ser Val Lys Ser
                325                 330                 335
Lys Val Gln Tyr Leu Lys Asp Arg Gln Leu Ala Gly Ala Met Val Trp
            340                 345                 350
Ala Leu Asp Leu Asp Asp Phe Gln Gly Ser Phe Cys Gly Gln Asp Leu
        355                 360                 365
Arg Phe Pro Leu Thr Asn Ala Ile Lys Asp Ala Leu Ala Ala Thr
    370                 375                 380
```

<210> 46
<211> 1528
<212> DNA
<213> Homo sapiens

<400> 46

```
ccggctccca ttccggctcc agcctccaat ccgaccccca tttcggctgc agcctcggac 60
ctagctccgg ccctcggtct atccggttgc atcctcctc cctgttccgg atcttatctt 120
gcgccacgcgc ctactccagg atcccgtagc cagacctcaa gccatggctg gtccttctc 180
ccgtctgctg tccgcccgcc cgggactcag gctcctggct ttggccggag cggggtctct 240
agccgctggg tttctgctcc gaccggaacc tgtacgagct gccatgaac gacggaggct 300
gtatccccog agcgctgagt acccagacct ccgaaagcac aacaactgca tggccagtca 360
cctgacccca gcagtctatg cacggctctg cgacaagacc acacccactg gttggacgct 420
agatcagtgt atccagactg gcgtggacaa ccctggccac cccttcatca agactgtggg 480
catggtggct ggagatgagg agacctatga ggtatttgct gacctgtttg acctgtgat 540
ccaagagcga cacaatggat atgacccccg gacaatgaag cacaccacgg atctagatgc 600
cagtaaaatc cgttctggct actttgatga gaggtatgta ttgtcctcta gagtcagaac 660
tggccgaagc atccgaggac tcagtctgcc tccagcttgc actcgagcag agcgacgaga 720
ggtggaacgt gttgtggtgg atgcactgag tggcctgaag ggtgacctgg ctggacgtta 780
ctataggctc agtgagatga cagaggctga acagcagcag cttattgatg accactttct 840
gtttgataag cctgtgtccc cgttgctgac tgcagcagga atggctccag actggccaga 900
tgctcgtgga atttggcaca acaatgagaa gagcttcctg atctgggtga atgaggagga 960
tcatacacgg gtgatctcca tggagaaggg tggtaacatg aagagagtgt ttgaaagatt 1020
ctgccgaggc ctcaaagagg tggagagact tatccaagaa cgtggctggg agttcatgtg 1080
gaatgagcgt ttgggataca tcttgacctg tccatctaac ctgggcactg gacttcgggc 1140
aggagtgcac atcaaactgc ccctgctaag caaagatagc cgcttcccaa agatcctgga 1200
gaacctaaga ctccaaaaac gtggtactgg aggagtggac actgctgcta caggcggtgt 1260
ctttgatatt tctaatttgg accgactagg caaatcagag gtggagctgg tgcaactggt 1320
catcgatgga gtaaactatt tgattgattg tgaacggcgt ctggagagag gccaggatat 1380
ccgcatcccc acacctgtca tccacaccaa gcattaactc cccatcgcca gctgatgact 1440
caagattccc aggagttttg ctcattctaa tgatggccca ttctacttgc tctggacctg 1500
ccccgcatc ccctgcctcc atcctagt                                    1528
```

<210> 47
<211> 417
<212> PRT
<213> Homo sapiens

<400> 47

```
Met Ala Gly Pro Phe Ser Arg Leu Leu Ser Ala Arg Pro Gly Leu Arg
 1               5                  10                  15
Leu Leu Ala Leu Ala Gly Ala Gly Ser Leu Ala Ala Gly Phe Leu Leu
            20                  25                  30
Arg Pro Glu Pro Val Arg Ala Ala Ser Glu Arg Arg Arg Leu Tyr Pro
        35                  40                  45
```

Pro Ser Ala Glu Tyr Pro Asp Leu Arg Lys His Asn Asn Cys Met Ala
50 55 60
Ser His Leu Thr Pro Ala Val Tyr Ala Arg Leu Cys Asp Lys Thr Thr
65 70 75 80
Pro Thr Gly Trp Thr Leu Asp Gln Cys Ile Gln Thr Gly Val Asp Asn
85 90 95
Pro Gly His Pro Phe Ile Lys Thr Val Gly Met Val Ala Gly Asp Glu
100 105 110
Glu Thr Tyr Glu Val Phe Ala Asp Leu Phe Asp Pro Val Ile Gln Glu
115 120 125
Arg His Asn Gly Tyr Asp Pro Arg Thr Met Lys His Thr Thr Asp Leu
130 135 140
Asp Ala Ser Lys Ile Arg Ser Gly Tyr Phe Asp Glu Arg Tyr Val Leu
145 150 155 160
Ser Ser Arg Val Arg Thr Gly Arg Ser Ile Arg Gly Leu Ser Leu Pro
165 170 175
Pro Ala Cys Thr Arg Ala Glu Arg Arg Glu Val Glu Arg Val Val Val
180 185 190
Asp Ala Leu Ser Gly Leu Lys Gly Asp Leu Ala Gly Arg Tyr Tyr Arg
195 200 205
Leu Ser Glu Met Thr Glu Ala Glu Gln Gln Gln Leu Ile Asp Asp His
210 215 220
Phe Leu Phe Asp Lys Pro Val Ser Pro Leu Leu Thr Ala Ala Gly Met
225 230 235 240
Ala Arg Asp Trp Pro Asp Ala Arg Gly Ile Trp His Asn Asn Glu Lys
245 250 255
Ser Phe Leu Ile Trp Val Asn Glu Glu Asp His Thr Arg Val Ile Ser
260 265 270
Met Glu Lys Gly Gly Asn Met Lys Arg Val Phe Glu Arg Phe Cys Arg
275 280 285
Gly Leu Lys Glu Val Glu Arg Leu Ile Gln Glu Arg Gly Trp Glu Phe
290 295 300
Met Trp Asn Glu Arg Leu Gly Tyr Ile Leu Thr Cys Pro Ser Asn Leu
305 310 315 320
Gly Thr Gly Leu Arg Ala Gly Val His Ile Lys Leu Pro Leu Leu Ser
325 330 335
Lys Asp Ser Arg Phe Pro Lys Ile Leu Glu Asn Leu Arg Leu Gln Lys
340 345 350
Arg Gly Thr Gly Gly Val Asp Thr Ala Ala Thr Gly Gly Val Phe Asp
355 360 365
Ile Ser Asn Leu Asp Arg Leu Gly Lys Ser Glu Val Glu Leu Val Gln
370 375 380
Leu Val Ile Asp Gly Val Asn Tyr Leu Ile Asp Cys Glu Arg Arg Leu
385 390 395 400
Glu Arg Gly Gln Asp Ile Arg Ile Pro Thr Pro Val Ile His Thr Lys
405 410 415
His

<210> 48
<211> 2365
<212> DNA
<213> Homo sapiens

<400> 48
ggagccggag agcgagcgcg gctgcagccg gcggcatggc tagcacggct tcggagatca 60
tcgccttcat ggtctccatc tcaggctggg tactggtgtc ctccacgctg cccaccgact 120

actggaaggt gtctaccatc gacggcacgg tcatcacaac cgccacctat tgggccaacc 180
tgtggaaggc gtgcgttacc gactccacgg gcgtctccaa ctgcaaggac ttcccctcca 240
tgctggcgct ggacggttat atacaggcat gtagaggact tatgatcgct gctgtcagcc 300
tgggcttctt tggttccata tttgcgctct ttggaatgaa gtgtaccaaa gtcggaggct 360
ccgataaagc caaagctaaa attgcttgtt tggctgggat tgtattcata ctgtcagggc 420
tgtgctcaat gactggatgt tccctatatg caaacaaaat cacaacggaa ttctttgatc 480
ctctctttgt tgagcaaaag tatgaattag gagccgctct gtttattgga tgggcaggag 540
cctcactgtg cataattggt ggtgtcatat tttgctttta aatatctgac aacaacaaaa 600
cacccagata cacatacaac ggggccacat ctgtcatgtc ttctcggaca aagtatcatg 660
gtggagaaga ttttaaaaca acaaaccctt caaaacagtt tgataaaaat gcttatgtct 720
aaaagagctc gctggcaagc tgcctcttga gtttgttata aaagcgaact gttcacaaaa 780
tgatccccatc aaggcctcc catcaattaac actcaaaact atttttaaaa tatgcatttg 840
aagcatctgt tgattgtatg gatgtaagtg ttcttacata gttagttata tactaatcat 900
tttctgttgt ggctttctat aaaaaataaa cagtttattt acaggatttg taaaatgttt 960
tctacattta tatagaacat gaaaagcatt tagtaccaaa ggttcaagaa gtattcgtac 1020
tctagccttt ttaatcattc atagatagaa gtctttgtac ccactcctta tgtttctttt 1080
cattcataaa caggtgtata aggaacaatg tcttataaac agcatgtggg caatctgaga 1140
atattcctca aaaggtgtcc aggttaaata gacatgttac tggctgcaca caggcaaatt 1200
ctagtttgtt ttttttaagt attctacaac atttatttaa aaaggtaaat ctttttgttg 1260
aagcagcaag ttatctggta gaacttaact tctacaggat cagagaggat cttgctcatt 1320
catggccata tccacatgcc catggccact cagtagattg ttgaaaaagc aaagccacac 1380
cattctcttt gatgtatgca gagagttacg tagcagggga tgttctctga ttctatccac 1440
tggcaccatt agtgaatatt tagttgtttt cataaacgat gctgtgatga agactcatgt 1500
acatatttag caaattttgg tttcttacat gtgcctgtca tgactgtaat tcattatgac 1560
tgctccagga agggctaatg gggccaatat attattgcct gtcatgtggc acatccatgt 1620
taaggggctg aggcgtccct ggcacggaat gcagagccct gagctagggc atcagcagaa 1680
gctgagatag agatattggt catggttgac tgaggagcca attaaaacct gtttatgcct 1740
agtgttccat tattggaaca ctaagcatgt gggagttatt tatatcctac tgctcaaggt 1800
catcgccaag gtgtgattgg aaaaattcaa aaaattgcaa cctcaggcat aaatgggtta 1860
aggacatccc aagcccaagt ggtacgtgcc tcactcagaa ctgacgggcc gagttctatc 1920
taggtgtgtc ttccagaacc tgtttacggc taactggata actgagagac ttgtcatttc 1980
taaagacatt taagttgctc cagggatttc tgaaaaaaga cacaggcttc ttcctagagc 2040
cagccctata taacatgccc acaagggcaa cagttatcac agttcataca caccttttcat 2100
gtcctgtctc actcactcct cacagccatc ctaggagata catattgttt tcatcctgca 2160
tttacagaaa aagaaatgaa aacagagagc ttaaataatt tgccacagta atgtcgaaac 2220
taggcctttg aaccaaggca gtctagggta aaatatagtt tcaaagtatg aataagaatt 2280
ggtatttgtg ttatctttga gtaagaaact gtccgatatg aatcacaacg tgggtgaatg 2340
tagtattttc ctgaagtgtg aaaga 2365

<210> 49
<211> 228
<212> PRT
<213> Homo sapiens

<400> 49
Met Ala Ser Thr Ala Ser Glu Ilc Ile Ala Phe Met Val Ser Ile Ser
1 5 10 15
Gly Trp Val Leu Val Ser Ser Thr Leu Pro Thr Asp Tyr Trp Lys Val
20 25 30
Ser Thr Ile Asp Gly Thr Val Ile Thr Thr Ala Thr Tyr Trp Ala Asn
35 40 45
Leu Trp Lys Ala Cys Val Thr Asp Ser Thr Gly Val Ser Asn Cys Lys
50 55 60
Asp Phe Pro Ser Met Leu Ala Leu Asp Gly Tyr Ile Gln Ala Cys Arg
65 70 75 80
Gly Leu Met Ile Ala Ala Val Ser Leu Gly Phe Phe Gly Ser Ile Phe
85 90 95
Ala Leu Phe Gly Met Lys Cys Thr Lys Val Gly Gly Ser Asp Lys Ala

```
            100                 105                 110
Lys Ala Lys Ile Ala Cys Leu Ala Gly Ile Val Phe Ile Leu Ser Gly
        115                 120                 125
Leu Cys Ser Met Thr Gly Cys Ser Leu Tyr Ala Asn Lys Ile Thr Thr
    130                 135                 140
Glu Phe Phe Asp Pro Leu Phe Val Glu Gln Lys Tyr Glu Leu Gly Ala
145                 150                 155                 160
Ala Leu Phe Ile Gly Trp Ala Gly Ala Ser Leu Cys Ile Ile Gly Gly
                165                 170                 175
Val Ile Phe Cys Phe Ser Ile Ser Asp Asn Asn Lys Thr Pro Arg Tyr
            180                 185                 190
Thr Tyr Asn Gly Ala Thr Ser Val Met Ser Ser Arg Thr Lys Tyr His
        195                 200                 205
Gly Gly Glu Asp Phe Lys Thr Thr Asn Pro Ser Lys Gln Phe Asp Lys
    210                 215                 220
Asn Ala Tyr Val
225
```

<210> 50
<211> 1024
<212> DNA
<213> Homo sapiens

<400> 50

```
ccccacccga aacacactca gcccttgcac tgacctgcct tctgattgga ggctggttgc 60
ttcggataat gacctccagg accccactgt tggttacagc ctgtttgtat tattcttact 120
gcaactcaag acacctgcag cagggcgtga gaaaaagtaa aagaccagta ttttcacatt 180
gccaggtacc agaaacacag aaqactgaca cccgccactt aagtggggcc agggctggtg 240
tctgcccatg ttgccatcct gatgggctgc ttgccacaat gagggatctt cttcaataca 300
tcgcttgctt ctttgccttt ttctctgctg ggtttttgat tgtggccacc tggactgact 360
gttggatggt gaatgctgat gactctctgg aggtgagcac aaaatgccga ggcctctggt 420
gggaatgcgt cacaaatgct tttgatggga ttcgcacctg tgatgagtac gattccatac 480
ttgcggagca tccttgaag ctggtggtaa ctcgagcgtt gatgattact gcagatattc 540
tagctgggtt tggattttct accctgctcc ttggtcttga ctgcgtgaaa ttcctccctg 600
atgagccgta cattaaagtc cgcatctgct ttgttgctgg agccacgtta ctaatagcag 660
gtaccccagg aatcattggc tctgtgtggt atgctgttga tgtgtatgtg gaacgttcta 720
ctttggtttt gcacaatata tttcttggta tccaatataa atttggttgg tcctgttggc 780
tcggaatggc tgggtctctg ggttgctttt tggctggagc tgttctcacc tgctgcttat 840
atcttttaa agatgttgga cctgagagaa actatcctta ttccttgagg aaagcctatt 900
cagccgcggg tgtttccatg gccaagtcat actcagcccc tcgcacagag acggccaaaa 960
tgtatgctgt agacacaagg gtgtaaaatg cacgtttcag ggtgtgtttg catatgattt 1020
aatc                                                          1024
```

<210> 51
<211> 305
<212> PRT
<213> Homo sapiens

<400> 51

```
Met Thr Ser Arg Thr Pro Leu Leu Val Thr Ala Cys Leu Tyr Tyr Ser
 1               5                  10                  15
Tyr Cys Asn Ser Arg His Leu Gln Gln Gly Val Arg Lys Ser Lys Arg
            20                  25                  30
Pro Val Phe Ser His Cys Gln Val Pro Glu Thr Gln Lys Thr Asp Thr
        35                  40                  45
Arg His Leu Ser Gly Ala Arg Ala Gly Val Cys Pro Cys Cys His Pro
    50                  55                  60
```

```
Asp Gly Leu Leu Ala Thr Met Arg Asp Leu Leu Gln Tyr Ile Ala Cys
65                  70                  75                  80
Phe Phe Ala Phe Phe Ser Ala Gly Phe Leu Ile Val Ala Thr Trp Thr
                85                  90                  95
Asp Cys Trp Met Val Asn Ala Asp Asp Ser Leu Glu Val Ser Thr Lys
            100                 105                 110
Cys Arg Gly Leu Trp Trp Glu Cys Val Thr Asn Ala Phe Asp Gly Ile
        115                 120                 125
Arg Thr Cys Asp Glu Tyr Asp Ser Ile Leu Ala Glu His Pro Leu Lys
    130                 135                 140
Leu Val Val Thr Arg Ala Leu Met Ile Thr Ala Asp Ile Leu Ala Gly
145                 150                 155                 160
Phe Gly Phe Leu Thr Leu Leu Leu Gly Leu Asp Cys Val Lys Phe Leu
                165                 170                 175
Pro Asp Glu Pro Tyr Ile Lys Val Arg Ile Cys Phe Val Ala Gly Ala
            180                 185                 190
Thr Leu Leu Ile Ala Gly Thr Pro Gly Ile Ile Gly Ser Val Trp Tyr
        195                 200                 205
Ala Val Asp Val Tyr Val Glu Arg Ser Thr Leu Val Leu His Asn Ile
    210                 215                 220
Phe Leu Gly Ile Gln Tyr Lys Phe Gly Trp Ser Cys Trp Leu Gly Met
225                 230                 235                 240
Ala Gly Ser Leu Gly Cys Phe Leu Ala Gly Ala Val Leu Thr Cys Cys
                245                 250                 255
Leu Tyr Leu Phe Lys Asp Val Gly Pro Glu Arg Asn Tyr Pro Tyr Ser
            260                 265                 270
Leu Arg Lys Ala Tyr Ser Ala Ala Gly Val Ser Met Ala Lys Ser Tyr
        275                 280                 285
Ser Ala Pro Arg Thr Glu Thr Ala Lys Met Tyr Ala Val Asp Thr Arg
    290                 295                 300
Val
305
```

<210> 52
<211> 1665
<212> DNA
<213> Homo sapiens

<400> 52

```
gaaggaactg gttctgctca cacttgctgg cttgcgcatc aggactggct ttatctcctg 60
actcacggtg caaaggtgca ctctgcgaac gttaagtccg tccccagcgc ttggaatcct 120
acggccccca cagccggatc ccctcagcct tccaggtcct caactcccgt ggacgctgaa 180
caatggcctc catgggggcta caggtaatgg gcatcgcgct ggccgtcctg ggctggctgg 240
ccgtcatgct gtgctgcgcg ctgcccatgt ggcgcgtgac ggccttcatc ggcagcaaca 300
ttgtcacctc gcagaccatc tgggagggcc tatggatgaa ctgcgtggtg cagagcaccg 360
gccagatgca gtgcaaggtg tacgactcgc tgctggcact gccgcaggac ctgcaggcgg 420
cccgcgccct cgtcatcatc agcatcatcg tggctgtctc gggcgtgctg ctgtccgtgg 480
tgggggggcaa gtgtaccaac tgcctggagg atgaaagcgc caaggccaag accatgatcg 540
tggcgggcgt ggtgttcctg ttggccggcc ttatggtgat agtgccggtg tcctggacgg 600
cccacaacat catccaagac ttctacaatc cgctggtggc ctccgggcag aagcgggaga 660
tgggtgcctc gctctacgtc ggctgggccg cctccggcct gctgctcctt ggcgggggge 720
tgctttgctg caactgtcca ccccgcacag acaagcctta ctccgccaag tattctgctg 780
cccgctctgc tgctgccagc aactacgtgt aaggtgccac ggctccactc tgttcctctc 840
tgctttgttc ttccctggac tgagctcagc gcaggtgtg accccaggag ggcctgcca 900
cgggccactg gctgctgggg actgggcagc tgagccaggc aggaaggcag 960
cagccttcag cctctctggc ccactcggac aacttcccaa ggccgctcc tgctagcaag 1020
aacagagtcc accctcctct ggatattggg gagggacgga agtgacaggg tgtggtggtg 1080
```

gagtggggag ctggcttctg ctggccagga tagcttaacc ctgactttgg gatctgcctg 1140
catcggcgtt ggccactgtc cccatttaca ttttccccac tctgtctgcc tgcatctcct 1200
ctgttccggg taggccttga tatcacctct gggactgtgc cttgctcacc gaaacccgcg 1260
cccaggagta tggctgaggc cttgcccacc cacctgcctg ggaagtgcag agtggatgga 1320
cgggtttaga ggggaggggc gaaggtgctg taaacaggtt tgggcagtgg tggggaggg 1380
ggccagagag gcggctcagg ttgcccagct ctgtggcctc aggactctct gcctcacccg 1440
cttcagccca gggcccctgg agactgatcc cctctgagtc ctctgcccct tccaaggaca 1500
ctaatgagcc tgggagggtg gcagggagga ggggacagct tcacccttgg aagtcctggg 1560
gtttttcctc ttccttcttt gtggtttctg ttttgtaatt taagaagagc tattcatcac 1620
tgtaattatt attattttct acaataaatg ggacctgtgc acagg 1665

<210> 53
<211> 209
<212> PRT
<213> Homo sapiens

<400> 53
Met Ala Ser Met Gly Leu Gln Val Met Gly Ile Ala Leu Ala Val Leu
1 5 10 15
Gly Trp Leu Ala Val Met Leu Cys Cys Ala Leu Pro Met Trp Arg Val
20 25 30
Thr Ala Phe Ile Gly Ser Asn Ile Val Thr Ser Gln Thr Ile Trp Glu
35 40 45
Gly Leu Trp Met Asn Cys Val Val Gln Ser Thr Gly Gln Met Gln Cys
50 55 60
Lys Val Tyr Asp Ser Leu Leu Ala Leu Pro Gln Asp Leu Gln Ala Ala
65 70 75 80
Arg Ala Leu Val Ile Ile Ser Ile Ile Val Ala Ala Leu Gly Val Leu
85 90 95
Leu Ser Val Val Gly Gly Lys Cys Thr Asn Cys Leu Glu Asp Glu Ser
100 105 110
Ala Lys Ala Lys Thr Met Ile Val Ala Gly Val Val Phe Leu Leu Ala
115 120 125
Gly Leu Met Val Ile Val Pro Val Ser Trp Thr Ala His Asn Ile Ile
130 135 140
Gln Asp Phe Tyr Asn Pro Leu Val Ala Ser Gly Gln Lys Arg Glu Met
145 150 155 160
Gly Ala Ser Leu Tyr Val Gly Trp Ala Ala Ser Gly Leu Leu Leu Leu
165 170 175
Gly Gly Gly Leu Leu Cys Cys Asn Cys Pro Pro Arg Thr Asp Lys Pro
180 185 190
Tyr Ser Ala Lys Tyr Ser Ala Ala Arg Ser Ala Ala Ala Ser Asn Tyr
195 200 205
Val

<210> 54
<211> 3457
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(3457)
<223> n = A,T,C or G

<400> 54

atgaagattt tgatacttgg tatttttctg tttttatgta gtaccccagc ctgggcgaaa 60
gaaaagcatt attacattgg aattattgaa acgacttggg attatgcctc tgaccatggg 120
gaaaagaaac ttatttctgt tgacacggaa cattccaata tctatcttca aaatggccca 180
gatagaattg ggagactata taagaaggcc ctttatcttc agtacacaga tgaaaccttt 240
aggacaacta tagaaaaacc ggtctggctt gggtttttag gccttattat caaagctgaa 300
actggagata aagtttatgt acacttaaaa aaccttgcct ctaggcccta cacctttcat 360
tcacatggaa taacttacta taaggaacat gaggggggca tctaccctga taacaccaca 420
gattttcaaa gagcagatga caaagtatat ccaggagagc agtatacata catgttgctt 480
gccactgaag aacaaagtcc tggggaagga gatggcaatt gtgtgactag gatttaccat 540
tcccacattg atgctccaaa agatattgcc tcaggactca tcggaccttt aataatctgt 600
aaaaaagatt ctctagataa agaaaaagaa aaacatattg accgagaatt tgtggtgatg 660
ttttctgtgg tggatgaaaa tttcagctgg tacctagaag acaacattaa aacctactgc 720
tcagaaccag agaagttgca caaagacaac gaagacttcc aggagagtaa cagaatgtat 780
tctgtgaatg gatacacttt tggaagtctc ccaggactct ccatgtgtgc tgaagacaga 840
gtaaaatggt acctttttgg tatgggtaat gaagttgatg tgcacgcagc tttcttttcac 900
gggcaagcac tgactaacaa gaactaccgt attgacacaa tcaacctctt tcctgctacc 960
ctgtttgatg cttatatggt ggcccagaac cctggagaat ggatgctcag ctgtcagaat 1020
ctaaaccatc tgaaagccgg tttgcaagcc tttttccagg tccaggagtg taacaagtct 1080
tcatcaaagg ataatatccg tgggaagcat gttagacact actacattgc cgctgaggaa 1140
atcatctgga actatgctcc ctctggtata gacatcttca ctaaagaaaa cttaacagca 1200
cctggaagtg actcagcggt gttttttgaa caaggtacca caagaattgg aggctcttat 1260
aaaaagctgg tttatcgtga gtacacagat gcctccttca caaatcgaaa ggagagaggc 1320
cctgaagaag agcatctttg catcctgggt cctgtcattt gggcagaggt gggagacacc 1380
atcagagtaa cctttcataa caaaggagca tatcccctca gtattggagcc gattgggtg 1440
agattcaata agaacaacga gggcacatac tattccccaa attacaaccc ccagagcaga 1500
agtgtgcctc cttcagcctc ccatgtggca cccacagaaa cattcaccta tgaatggact 1560
gtccccaaag aagtaggacc cactaatgca gatcctgtgt gtctagctaa gatgtattat 1620
tctgctgtgg atcccactaa agatatattc actgggctta ttgggccaat gaaaatatgc 1680
aagaaaggaa gtttacatgc aaatgggaga cagaaagatg tagacaagga attctatttg 1740
tttcctacag tatttgatga gaatgagagt ttactcctgg aagataatat tagaatgttt 1800
acaactgcac ctgatcaggt ggataaggaa gatgaagact ttcaggaatc taataaaatg 1860
cactccatga atggattcat gtatgggaat cagccgggtc tcactatgtg caaaggagat 1920
tcggtcgtgt ggtacttatt cagcgccgga aatgaggccg atgtacatgg aatatacttt 1980
tcaggaaaca catatctgtg gagaggagaa cggagagaca cagcaaacct cttccctcaa 2040
acaagtctta cgctccacat gtggcctgac acagagggga cttttaatgt tgaatgcctt 2100
acaactgatc attacacagg cggcatgaag caaaaatata ctgtgaacca atgcaggcgg 2160
cagtctgagg attccacctt ctacctggga gagaggacat actatatcgc agcagtggag 2220
gtggaatggg attattcccc acaaagggag tgggaaaagg agctgcatca ttacaagagg 2280
cagaatgttt caaatgcatt tttagataag ggagagtttt acatagggctc aaagtacaag 2340
aaagttgtgt atcggcagta tactgatagc acattccgtg ttccagtgga gagaaagct 2400
gaagaagaac atctgggaat tctaggtcca caacttcatg cagatgttgg agacaaagtc 2460
aaaattatct ttaaaaacat ggccacaagg ccctactcaa tacatgccca tggggtacaa 2520
acagagagtt ctacagttac tccaacatta ccaggtgaaa ctctcactta cgtatggaaa 2580
atcccagaaa gattctggagc tggaacagag gattctgctt gtattccatg ggctttattat 2640
tcaactgtgg atcaagttaa ggacctctac agtggattaa ttggcccct gattgtttgt 2700
cgaagacctt acttgaaagt attcaatccc agaaggaagc tggaatttgc ccttctgttt 2760
ctagtttttg atgagaatga atcttggtac ttagatgaca acatcaaaac atactctgat 2820
caccccgaga aagtaaacaa agatgatgag gaattcatag aaagcaataa agtcatgct 2880
attaatggaa gaatgtttgg aaacctacaa ggcctcacaa tgcacgtggg agatgaagtc 2940
aactggtatc tgatgggaat gggcaatgaa atagacttac acactgtaca ttttcacggc 3000
catagcttcc aatacaagca caggggagtt tatagttctg atgtctttga catttccct 3060
ggaacatacc aaccctaga aatgtttcca agaacacctg gaatttggtt actccactgc 3120
catgtgaccg accacattca tgctggaatg gaaaccactt acaccgttct acaaaatgaa 3180
ggtgaatatc cagcaccaa atctggctga atgaaatgaa ttggtgataa gtggaaaaaa 3240
gagaaaaacc aatgattcat aacaatgtat gtgaaagtgt aaaatagaat gttactttgg 3300
aatgactata aacattaaaa gnagactggn agcatacaac tttgtacatt tgtgggggaa 3360
aactattaat tttgcccaaa tggaaagatc aacagactat ataatgatac atgactgaca 3420
cttgtacact aggtaataaa actgattcat acagtct 3457

<210> 55
<211> 1069
<212> PRT
<213> Homo sapiens

<400> 55
Met Lys Ile Leu Ile Leu Gly Ile Phe Leu Phe Leu Cys Ser Thr Pro
1 5 10 15
Ala Trp Ala Lys Glu Lys His Tyr Tyr Ile Gly Ile Ile Glu Thr Thr
20 25 30
Trp Asp Tyr Ala Ser Asp His Gly Glu Lys Lys Leu Ile Ser Val Asp
35 40 45
Thr Glu His Ser Asn Ile Tyr Leu Gln Asn Gly Pro Asp Arg Ile Gly
50 55 60
Arg Leu Tyr Lys Lys Ala Leu Tyr Leu Gln Tyr Thr Asp Glu Thr Phe
65 70 75 80
Arg Thr Thr Ile Glu Lys Pro Val Trp Leu Gly Phe Leu Gly Pro Ile
85 90 95
Ile Lys Ala Glu Thr Gly Asp Lys Val Tyr Val His Leu Lys Asn Leu
100 105 110
Ala Ser Arg Pro Tyr Thr Phe His Ser His Gly Ile Thr Tyr Tyr Lys
115 120 125
Glu His Glu Gly Ala Ile Tyr Pro Asp Asn Thr Thr Asp Phe Gln Arg
130 135 140
Ala Asp Asp Lys Val Tyr Pro Gly Glu Gln Tyr Thr Tyr Met Leu Leu
145 150 155 160
Ala Thr Glu Glu Gln Ser Pro Gly Glu Gly Asp Gly Asn Cys Val Thr
165 170 175
Arg Ile Tyr His Ser His Ile Asp Ala Pro Lys Asp Ile Ala Ser Gly
180 185 190
Leu Ile Gly Pro Leu Ile Ile Cys Lys Lys Asp Ser Leu Asp Lys Glu
195 200 205
Lys Glu Lys His Ile Asp Arg Glu Phe Val Val Met Phe Ser Val Val
210 215 220
Asp Glu Asn Phe Ser Trp Tyr Leu Glu Asp Asn Ile Lys Thr Tyr Cys
225 230 235 240
Ser Glu Pro Glu Lys Val Asp Lys Asp Asn Glu Asp Phe Gln Glu Ser
245 250 255
Asn Arg Met Tyr Ser Val Asn Gly Tyr Thr Phe Gly Ser Leu Pro Gly
260 265 270
Leu Ser Met Cys Ala Glu Asp Arg Val Lys Trp Tyr Leu Phe Gly Met
275 280 285
Gly Asn Glu Val Asp Val His Ala Ala Phe Phe His Gly Gln Ala Leu
290 295 300
Thr Asn Lys Asn Tyr Arg Ile Asp Thr Ile Asn Leu Phe Pro Ala Thr
305 310 315 320
Leu Phe Asp Ala Tyr Met Val Ala Gln Asn Pro Gly Glu Trp Met Leu
325 330 335
Ser Cys Gln Asn Leu Asn His Leu Lys Ala Gly Leu Gln Ala Phe Phe
340 345 350
Gln Val Gln Glu Cys Asn Lys Ser Ser Ser Lys Asp Asn Ile Arg Gly
355 360 365
Lys His Val Arg His Tyr Tyr Ile Ala Ala Glu Glu Ile Ile Trp Asn
370 375 380
Tyr Ala Pro Ser Gly Ile Asp Ile Phe Thr Lys Glu Asn Leu Thr Ala
385 390 395 400
Pro Gly Ser Asp Ser Ala Val Phe Phe Glu Gln Gly Thr Thr Arg Ile

```
                405                 410                 415
Gly Gly Ser Tyr Lys Lys Leu Val Tyr Arg Glu Tyr Thr Asp Ala Ser
            420                 425                 430
Phe Thr Asn Arg Lys Glu Arg Gly Pro Glu Glu Glu His Leu Gly Ile
        435                 440                 445
Leu Gly Pro Val Ile Trp Ala Glu Val Gly Asp Thr Ile Arg Val Thr
    450                 455                 460
Phe His Asn Lys Gly Ala Tyr Pro Leu Ser Ile Glu Pro Ile Gly Val
465                 470                 475                 480
Arg Phe Asn Lys Asn Asn Glu Gly Thr Tyr Tyr Ser Pro Asn Tyr Asn
                485                 490                 495
Pro Gln Ser Arg Ser Val Pro Pro Ser Ala Ser His Val Ala Pro Thr
            500                 505                 510
Glu Thr Phe Thr Tyr Glu Trp Thr Val Pro Lys Glu Val Gly Pro Thr
        515                 520                 525
Asn Ala Asp Pro Val Cys Leu Ala Lys Met Tyr Tyr Ser Ala Val Asp
    530                 535                 540
Pro Thr Lys Asp Ile Phe Thr Gly Leu Ile Gly Pro Met Lys Ile Cys
545                 550                 555                 560
Lys Lys Gly Ser Leu His Ala Asn Gly Arg Gln Lys Asp Val Asp Lys
                565                 570                 575
Glu Phe Tyr Leu Phe Pro Thr Val Phe Asp Glu Asn Glu Ser Leu Leu
            580                 585                 590
Leu Glu Asp Asn Ile Arg Met Phe Thr Thr Ala Pro Asp Gln Val Asp
        595                 600                 605
Lys Glu Asp Glu Asp Phe Gln Glu Ser Asn Lys Met His Ser Met Asn
    610                 615                 620
Gly Phe Met Tyr Gly Asn Gln Pro Gly Leu Thr Met Cys Lys Gly Asp
625                 630                 635                 640
Ser Val Val Trp Tyr Leu Phe Ser Ala Gly Asn Glu Ala Asp Val His
                645                 650                 655
Gly Ile Tyr Phe Ser Gly Asn Thr Tyr Leu Trp Arg Gly Glu Arg Arg
            660                 665                 670
Asp Thr Ala Asn Leu Phe Pro Gln Thr Ser Leu Thr Leu His Met Trp
        675                 680                 685
Pro Asp Thr Glu Gly Thr Phe Asn Val Glu Cys Leu Thr Thr Asp His
    690                 695                 700
Tyr Thr Gly Gly Met Lys Gln Lys Tyr Thr Val Asn Gln Cys Arg Arg
705                 710                 715                 720
Gln Ser Glu Asp Ser Thr Phe Tyr Leu Gly Glu Arg Thr Tyr Tyr Ile
                725                 730                 735
Ala Ala Val Glu Val Glu Trp Asp Tyr Ser Pro Gln Arg Glu Trp Glu
            740                 745                 750
Lys Glu Leu His His Leu Gln Glu Gln Asn Val Ser Asn Ala Phe Leu
        755                 760                 765
Asp Lys Gly Glu Phe Tyr Ile Gly Ser Lys Tyr Lys Lys Val Val Tyr
    770                 775                 780
Arg Gln Tyr Thr Asp Ser Thr Phe Arg Val Pro Val Glu Arg Lys Ala
785                 790                 795                 800
Glu Glu Glu His Leu Gly Ile Leu Gly Pro Gln Leu His Ala Asp Val
                805                 810                 815
Gly Asp Lys Val Lys Ile Ile Phe Lys Asn Met Ala Thr Arg Pro Tyr
            820                 825                 830
Ser Ile His Ala His Gly Val Gln Thr Glu Ser Ser Thr Val Thr Pro
        835                 840                 845
Thr Leu Pro Gly Glu Thr Leu Thr Tyr Val Trp Lys Ile Pro Glu Arg
    850                 855                 860
Ser Gly Ala Gly Thr Glu Asp Ser Ala Cys Ile Pro Trp Ala Tyr Tyr
```

```
865                 870                 875                 880
Ser Thr Val Asp Gln Val Lys Asp Leu Tyr Ser Gly Leu Ile Gly Pro
                885                 890                 895
Leu Ile Val Cys Arg Arg Pro Tyr Leu Lys Val Phe Asn Pro Arg Arg
            900                 905                 910
Lys Leu Glu Phe Ala Leu Leu Phe Leu Val Phe Asp Glu Asn Glu Ser
        915                 920                 925
Trp Tyr Leu Asp Asp Asn Ile Lys Thr Tyr Ser Asp His Pro Glu Lys
    930                 935                 940
Val Asn Lys Asp Asp Glu Glu Phe Ile Glu Ser Asn Lys Met His Ala
945                 950                 955                 960
Ile Asn Gly Arg Met Phe Gly Asn Leu Gln Gly Leu Thr Met His Val
                965                 970                 975
Gly Asp Glu Val Asn Trp Tyr Leu Met Gly Met Gly Asn Glu Ile Asp
            980                 985                 990
Leu His Thr Val His Phe His Gly His Ser Phe Gln Tyr Lys His Arg
        995                 1000                1005
Gly Val Tyr Ser Ser Asp Val Phe Asp Ile Phe Pro Gly Thr Tyr Gln
    1010                1015                1020
Thr Leu Glu Met Phe Pro Arg Thr Pro Gly Ile Trp Leu Leu His Cys
1025                1030                1035                1040
His Val Thr Asp His Ile His Ala Gly Met Glu Thr Thr Tyr Thr Val
                1045                1050                1055
Leu Gln Asn Glu Gly Glu Tyr Pro Asp Thr Lys Ser Gly
            1060                1065
```

<210> 56
<211> 2807
<212> DNA
<213> Homo sapiens

<400> 56

```
cgtggtgatg ttttctgtgg tggatgaaaa tttcagctgg tacctagaag acaacattaa 60
aacctactgc tcagaaccag agaaagttga caaagacaac gaagacttcc aggagagtaa 120
cagaatgtat tctgtgaatg gatacacttt tggaagtctc ccaggactct ccatgtgtgc 180
tgaagacaga gtaaaatggt acctttttgg tatgggtaat gaagttgatg tgcacgcagc 240
tttctttcac gggcaagcac tgactaacaa gaactaccgt attgacacaa tcaacctctt 300
tcctgctacc ctgttttgatg cttatatggt ggcccagaac cctggagaat ggatgctcag 360
ctgtcagaat ctaaaccatc tgaaagccgg tttgcaagcc tcttttccagg tccaggagtg 420
taacaagtct tcatcaaagg ataatatccg tgggaagcat gttagacact actacattgc 480
cgctgaggaa atcatctgga actatgctcc ctctggtata gacatcttca ctaaagaaaa 540
cttaacagca cctggaagtg actcagcggt gtttttgaa caaggtacca caagaattgg 600
aggctcttat aaaaagctgg tttatcgtga gtacacagat gcctccttca caaatcgaaa 660
ggagagaggc cctgaagaag agcatcttgg catcctggat cctgtcattt gggcagaggt 720
gggagacacc atcagagtaa ccttccataa caaaggagca tatcccctca gtattgagcc 780
gattggggtg agattcaata agaacaacga gggcacatac tattccccaa attacaaccc 840
ccagagcaga agtgtgcctc cttcagcctc ccatgtggca cccacagaaa cattcaccta 900
tgaatggact gtcccaaag aagtaggacc cactaatgca gatcctgtgt gtctagctaa 960
gatgtattat tctgctgtgg atcccactaa agatatattc actggcttta ttgggccaat 1020
gaaaatatgc aagaaaggaa gtttacatgc aaatgggaga cagaaagatg tagacaagga 1080
attctatttg tttcctacag tatttgatga gaatgagagt ttactcctgg aagatataat 1140
tagaatgttt acaactgcac ctgatcaggt ggataaggaa gatgaagact ttcaggaatc 1200
taataaaatg cactccatga atggattcat gtatgggaat cagccgggtc tcactatgtg 1260
caaaggagat tcggtcgtgt ggtacttatt cagcgccgga aatgaggccg atgtacatgg 1320
aatatacttt tcaggaaaca catatctgtg gagaggagaa cggagagaca cagcaaacct 1380
cttcctcaa acaagtctta cgctccacat gtggcctgac acagagggga cttttaatgt 1440
tgaatgcctt acaactgatc attacacagg cggcatgaag caaaaatata ctgtgaacca 1500
```

atgcaggcgg cagtctgagg attccacctt ctacctggga gagaggacat actatatcgc 1560
agcagtggag gtggaatggg attattcccc acaaagggag tgggaaaagg agctgcatca 1620
tttacaagag cagaatgttt caaatgcatt tttagataag ggagagtttt acatagggctc 1680
aaagtacaag aaagttgtgt atcggcagta tactgatagc acattccgtg ttccagtgga 1740
gagaaaagct gaagaagaac atctgggaat tctaggtcca caacttcatg cagatgttgg 1800
agacaaagtc aaaattatct ttaaaaacat ggccacaagg ccctactcaa tacatgccca 1860
tggggtacaa acagagagtt ctacagttac tccaacatta ccaggtgaaa ctctcactta 1920
cgtatggaaa atcccagaaa gatctggagc tggaacagag gattctgctt gtattccatg 1980
ggcttattat tcaactgtgg atcaagttaa ggacctctac agtggattaa ttggcccccct 2040
gattgtttgt cgaagacctt acttgaaagt attcaatccc agaaggaaac tggaatttgc 2100
ccttctgttt ctagtttttg atgagaatga atcttggtac ttagatgaca acatcaaaac 2160
atacttctgat caccccgaga aagtaaacaa agatgatgag gaattcatag aaagcaataa 2220
aatgcatgct attaatggaa gaatgtttgg aaacctacaa ggcctcacaa tgcacgtggg 2280
agatgaagtc aactggtatc tgatgggaat gggcaatgaa atagacttac acactgtaca 2340
ttttcacggc catagcttcc aatacaagca cagggagtt tatagttctg atgtctttga 2400
cattttccct ggaacatacc aaaccctaga aatgtttcca agaacacctg gaatttggtt 2460
actccactgc catgtgaccg accacattca tgctggaatg gaaaccactt acaccgttct 2520
acaaaatgaa ggtgaatatc cagacaccaa atctggctga atgaaataaa ttggtgataa 2580
gtggaaaaaa gagaaaaacc aatgattcat aacaatgtat gtgaaagtgt aaaatagaat 2640
gttacttttgg aatgactata aacattaaaa gaagactgga agcatacaac tttgtacatt 2700
tgtgggggaa aactattaat tttgcccaaa tggaaagatc aacagactat ataatgatac 2760
atgactgaca cttgtacact aggtaataaa actgattcat acagtct 2807

<210> 57
<211> 852
<212> PRT
<213> Homo sapiens

<400> 57
Val Val Met Phe Ser Val Val Asp Glu Asn Phe Ser Trp Tyr Leu Glu
1 5 10 15
Asp Asn Ile Lys Thr Tyr Cys Ser Glu Pro Glu Lys Val Asp Lys Asp
20 25 30
Asn Glu Asp Phe Gln Glu Ser Asn Arg Met Tyr Ser Val Asn Gly Tyr
35 40 45
Thr Phe Gly Ser Leu Pro Gly Leu Ser Met Cys Ala Glu Asp Arg Val
50 55 60
Lys Trp Tyr Leu Phe Gly Met Gly Asn Glu Val Asp Val His Ala Ala
65 70 75 80
Phe Phe His Gly Gln Ala Leu Thr Asn Lys Asn Tyr Arg Ile Asp Thr
85 90 95
Ile Asn Leu Phe Pro Ala Thr Leu Phe Asp Ala Tyr Met Val Ala Gln
100 105 110
Asn Pro Gly Glu Trp Met Leu Ser Cys Gln Asn Leu Asn His Leu Lys
115 120 125
Ala Gly Leu Gln Ala Phe Phe Gln Val Gln Glu Cys Asn Lys Ser Ser
130 135 140
Ser Lys Asp Asn Ile Arg Gly Lys His Val Arg His Tyr Tyr Ile Ala
145 150 155 160
Ala Glu Glu Ile Ile Trp Asn Tyr Ala Pro Ser Gly Ile Asp Ile Phe
165 170 175
Thr Lys Glu Asn Leu Thr Ala Pro Gly Ser Asp Ser Ala Val Phe Phe
180 185 190
Glu Gln Gly Thr Thr Arg Ile Gly Gly Ser Tyr Lys Lys Leu Val Tyr
195 200 205
Arg Glu Tyr Thr Asp Ala Ser Phe Thr Asn Arg Lys Glu Arg Gly Pro
210 215 220
Glu Glu Glu His Leu Gly Ile Leu Gly Pro Val Ile Trp Ala Glu Val

```
225                 230                 235                 240
Gly Asp Thr Ile Arg Val Thr Phe His Asn Lys Gly Ala Tyr Pro Leu
                245                 250                 255
Ser Ile Glu Pro Ile Gly Val Arg Phe Asn Lys Asn Asn Glu Gly Thr
            260                 265                 270
Tyr Tyr Ser Pro Asn Tyr Asn Pro Gln Ser Arg Ser Val Pro Pro Ser
        275                 280                 285
Ala Ser His Val Ala Pro Thr Glu Thr Phe Thr Tyr Glu Trp Thr Val
    290                 295                 300
Pro Lys Glu Val Gly Pro Thr Asn Ala Asp Pro Val Cys Leu Ala Lys
305                 310                 315                 320
Met Tyr Tyr Ser Ala Val Asp Pro Thr Lys Asp Ile Phe Thr Gly Leu
                325                 330                 335
Ile Gly Pro Met Lys Ile Cys Lys Lys Gly Ser Leu His Ala Asn Gly
            340                 345                 350
Arg Gln Lys Asp Val Asp Lys Glu Phe Tyr Leu Phe Pro Thr Val Phe
        355                 360                 365
Asp Glu Asn Glu Ser Leu Leu Leu Glu Asp Asn Ile Arg Met Phe Thr
    370                 375                 380
Thr Ala Pro Asp Gln Val Asp Lys Glu Asp Glu Asp Phe Gln Glu Ser
385                 390                 395                 400
Asn Lys Met His Ser Met Asn Gly Phe Met Tyr Gly Asn Gln Pro Gly
                405                 410                 415
Leu Thr Met Cys Lys Gly Asp Ser Val Val Trp Tyr Leu Phe Ser Ala
            420                 425                 430
Gly Asn Glu Ala Asp Val His Gly Ile Tyr Phe Ser Gly Asn Thr Tyr
        435                 440                 445
Leu Trp Arg Gly Glu Arg Arg Asp Thr Ala Asn Leu Phe Pro Gln Thr
    450                 455                 460
Ser Leu Thr Leu His Met Trp Pro Asp Thr Glu Gly Thr Phe Asn Val
465                 470                 475                 480
Glu Cys Leu Thr Thr Asp His Tyr Thr Gly Gly Met Lys Gln Lys Tyr
                485                 490                 495
Thr Val Asn Gln Cys Arg Arg Gln Ser Glu Asp Ser Thr Phe Tyr Leu
            500                 505                 510
Gly Glu Arg Thr Tyr Tyr Ile Ala Ala Val Glu Val Glu Trp Asp Tyr
        515                 520                 525
Ser Pro Gln Arg Glu Trp Glu Lys Glu Leu His His Leu Gln Glu Gln
    530                 535                 540
Asn Val Ser Asn Ala Phe Leu Asp Lys Gly Glu Phe Tyr Ile Gly Ser
545                 550                 555                 560
Lys Tyr Lys Lys Val Val Tyr Arg Gln Tyr Thr Asp Ser Thr Phe Arg
                565                 570                 575
Val Pro Val Glu Arg Lys Ala Glu Glu Glu His Leu Gly Ile Leu Gly
            580                 585                 590
Pro Gln Leu His Ala Asp Val Gly Asp Lys Val Lys Ile Ile Phe Lys
        595                 600                 605
Asn Met Ala Thr Arg Pro Tyr Ser Ile His Ala His Gly Val Gln Thr
    610                 615                 620
Glu Ser Ser Thr Val Thr Pro Thr Leu Pro Gly Glu Thr Leu Thr Tyr
625                 630                 635                 640
Val Trp Lys Ile Pro Glu Arg Ser Gly Ala Gly Thr Glu Asp Ser Ala
                645                 650                 655
Cys Ile Pro Trp Ala Tyr Tyr Ser Thr Val Asp Gln Val Lys Asp Leu
            660                 665                 670
Tyr Ser Gly Leu Ile Gly Pro Leu Ile Val Cys Arg Arg Pro Tyr Leu
        675                 680                 685
Lys Val Phe Asn Pro Arg Arg Lys Leu Glu Phe Ala Leu Leu Phe Leu
```

```
    690                 695                 700
Val Phe Asp Glu Asn Glu Ser Trp Tyr Leu Asp Asp Asn Ile Lys Thr
705                 710                 715                 720
Tyr Ser Asp His Pro Glu Lys Val Asn Lys Asp Asp Glu Glu Phe Ile
                725                 730                 735
Glu Ser Asn Lys Met His Ala Ile Asn Gly Arg Met Phe Gly Asn Leu
            740                 745                 750
Gln Gly Leu Thr Met His Val Gly Asp Glu Val Asn Trp Tyr Leu Met
        755                 760                 765
Gly Met Gly Asn Glu Ile Asp Leu His Thr Val His Phe His Gly His
    770                 775                 780
Ser Phe Gln Tyr Lys His Arg Gly Val Tyr Ser Ser Asp Val Phe Asp
785                 790                 795                 800
Ile Phe Pro Gly Thr Tyr Gln Thr Leu Glu Met Phe Pro Arg Thr Pro
                805                 810                 815
Gly Ile Trp Leu Leu His Cys His Val Thr Asp His Ile His Ala Gly
            820                 825                 830
Met Glu Thr Thr Tyr Thr Val Leu Gln Asn Glu Gly Glu Tyr Pro Asp
        835                 840                 845
Thr Lys Ser Gly
    850
```

<210> 58
<211> 3321
<212> DNA
<213> Homo sapiens

<400> 58

```
atgaagattt tgatacttgg tatttttctg tttttatgta gtaccccagc ctgggcgaaa 60
gaaaagcatt attacattgg aattattgaa acgacttggg attatgcctc tgaccatggg 120
gaaaagaaac ttatttctgt tgacacggaa cattccaata tctatcttca aaatggccca 180
gatagaattg ggagactata taagaaggcc ctttatcttc agtacacaga tgaaacctt 240
aggacaacta tagaaaaacc ggtctggctt gggtttttag gcacctattat caaagctgaa 300
actggagata aagtttatgt acacttaaaa aaccttgcct ctaggcccta caccttttcat 360
tcacatggaa taacttacta taaggaacat gaggggggcca tctaccctga taacaccaca 420
gattttcaaa gagcagatga caaagtatat ccaggagagc agtatacata catgttgctt 480
gccactgaag aacaaagtcc tggggaagga gatggcaatt gtgtgactag gattaccat 540
tcccacattg atgctccaaa agatattgcc tcaggactca tcggacctt aataatctgt 600
aaaaaagatt ctctagataa agaaaaagaa aaacatattg accgagaatt tgtggtgatg 660
ttttctgtgg tggatgaaaa tttcagctgg tacctagaag acaacattaa aacctactgc 720
tcagaaccag agaaagttga caaagacaac gaagacttcc aggagagtaa cagaatgtat 780
tctgtgaatg gatacacttt tggaagtctc ccaggactct ccatgtgtgc tgaagacaga 840
gtaaaatggt acttttttgg tatgggtaat gaagttgatg tgcacgcagc tttctttcac 900
gggcaagcac tgactaacaa gaactaccgt attgacacaa tcaaccttct tcctgctacc 960
ctgtttgatg cttatatggt ggcccagaac cctggagaat ggatgctcag ctgtcagaat 1020
ctaaaccatc tgaaagccgg tttgcaagcc tttttccagg tccaggagtg taacaagtct 1080
tcatcaaagg ataatatccg tgggaagcat gttagacact actacattgc cgctgaggaa 1140
atcatctgga actatgctcc ctctggtata gacatcttca ctaaagaaaa cttaacagca 1200
cctggaagtg actcagcggt gtttttgaa caaggtacca caagaattgg aggctcttat 1260
aaaaagctgg tttatcgtga gtacacagat gcctccttca caaatcgaaa ggagagaggc 1320
cctgaagaag agcatctttgg catcctgggt cctgtcattt gggcagaggt ggagacaccc 1380
atcagagtaa ccttccataa caaaggagca tatccccaca gtattgagcc gattgggtg 1440
agattcaata agaacaacga gggcacatac tattccccaa attacaaccc ccagagcaga 1500
agtgtgcctc cttcagcctc ccatgtggca cccacagaaa cattcaccta tgaatggact 1560
gtccccaaag aagtaggacc cactaatgca gatcctgtgt gtctagctaa gatgtattat 1620
tctgctgtgg atcccactaa agatatattc actgggctta ttgggccaat gaaaatatgc 1680
aagaaaggaa gtttacatgc aaatgggaga cagaaagatg tagacaagga attctatttg 1740
```

tttcctacag tatttgatga gaatgagagt ttactcctgg aagataatat tagaatgttt 1800
acaactgcac ctgatcaggt ggataaggaa gatgaagact ttcaggaatc taataaaatg 1860
cactccatga atggattcat gtatgggaat cagccgggtc tcactatgtg caaaggagat 1920
tcggtcgtgt ggtacttatt cagcgccgga aatgaggccg atgtacatgg aatatacttt 1980
tcaggaaaca catatctgtg gagaggagaa cggagagaca cagcaaacct cttccctcaa 2040
acaagtctta cgctccacat gtggcctgac acagagggga cttttaatgt tgaatgcctt 2100
acaactgatc attacacagg cggcatgaag caaaaatata ctgtgaacca atgcaggcgg 2160
cagtctgagg attccacctt ctacctggga gagaggacat actatatcgc agcagtggag 2220
gtggaatggg attattcccc acaaagggag tgggaaaagg agctgcatca tttacaagag 2280
cagaatgttt caaatgcatt tttagataag ggagagtttt acatagggctc aaagtacaag 2340
aaagttgtgt atcggcagta tactgatagc acattccgtg ttccagtgga gagaaaagct 2400
gaagaagaac atctgggaat tctaggtcca caacttcatg cagatgttgg agacaaagtc 2460
aaaattatct ttaaaaacat ggccacaagg ccctactcaa tacatgccca tggggtacaa 2520
acagagagtt ctacagttac tccaacatta ccaggtgaaa ctctcactta cgtatggaaa 2580
atcccagaaa gatctggagc tggaacagag gattctgctt gtattccatg ggcttattat 2640
tcaactgtgg atcaagttaa ggacctctac agtggattaa ttggcccctt gattgtttgt 2700
cgaagacctt acttgaaagt attcaatccc agaaggaagc tggaatttgc ccttctgttt 2760
ctagtttttg atgagaatga atcttggtac ttagatgaca acatcaaaac atactctgat 2820
cacccgaga aagtaaacaa agatgatgag gaattcatag aaagcaataa aatgcatgct 2880
attaatggaa gaatgtttgg aaacctacaa ggcctcacaa tgcacgtggg agatgaagtc 2940
aactggtatc tgatgggaat gggcaatgaa atagacttac acactgtaca ttttcacggc 3000
catagcttcc aatacaagca caggggagtt tatagttctg atgtctttga cattttccct 3060
ggaacatacc aaaccctaga aatgtttcca agaacacctg gaatttgtgt actccactgc 3120
catgtgaccg accacattca tgctggaatg gaaaccactt acaccgttct acaaaatgaa 3180
gacaccaaat ctggctgaat gaaataaatt ggtgataagt ggaaaaaaga gaaaaaccaa 3240
tgattcataa caatgtatgt gaaagtgtaa aatagaatgt tactttggaa tgactataaa 3300
cattaaaaga gactggagca t 3321

<210> 59
<211> 1065
<212> PRT
<213> Homo sapiens

<400> 59
Met Lys Ile Leu Ile Leu Gly Ile Phe Leu Phe Leu Cys Ser Thr Pro
1 5 10 15
Ala Trp Ala Lys Glu Lys His Tyr Tyr Ile Gly Ile Ile Glu Thr Thr
20 25 30
Trp Asp Tyr Ala Ser Asp His Gly Glu Lys Lys Leu Ile Ser Val Asp
35 40 45
Thr Glu His Ser Asn Ile Tyr Leu Gln Asn Gly Pro Asp Arg Ile Gly
50 55 60
Arg Leu Tyr Lys Lys Ala Leu Tyr Leu Gln Tyr Thr Asp Glu Thr Phe
65 70 75 80
Arg Thr Thr Ile Glu Lys Pro Val Trp Leu Gly Phe Leu Gly Pro Ile
85 90 95
Ile Lys Ala Glu Thr Gly Asp Lys Val Tyr Val His Leu Lys Asn Leu
100 105 110
Ala Ser Arg Pro Tyr Thr Phe His Ser His Gly Ile Thr Tyr Tyr Lys
115 120 125
Glu His Glu Gly Ala Ile Tyr Pro Asp Asn Thr Thr Asp Phe Gln Arg
130 135 140
Ala Asp Asp Lys Val Tyr Pro Gly Glu Gln Tyr Thr Tyr Met Leu Leu
145 150 155 160
Ala Thr Glu Glu Gln Ser Pro Gly Glu Gly Asp Gly Asn Cys Val Thr
165 170 175
Arg Ile Tyr His Ser His Ile Asp Ala Pro Lys Asp Ile Ala Ser Gly
180 185 190

```
Leu Ile Gly Pro Leu Ile Ile Cys Lys Lys Asp Ser Leu Asp Lys Glu
        195                 200                 205
Lys Glu Lys His Ile Asp Arg Glu Phe Val Val Met Phe Ser Val Val
    210                 215                 220
Asp Glu Asn Phe Ser Trp Tyr Leu Glu Asp Asn Ile Lys Thr Tyr Cys
225                 230                 235                 240
Ser Glu Pro Glu Lys Val Asp Lys Asp Asn Glu Asp Phe Gln Glu Ser
                245                 250                 255
Asn Arg Met Tyr Ser Val Asn Gly Tyr Thr Phe Gly Ser Leu Pro Gly
            260                 265                 270
Leu Ser Met Cys Ala Glu Asp Arg Val Lys Trp Tyr Leu Phe Gly Met
        275                 280                 285
Gly Asn Glu Val Asp Val His Ala Ala Phe Phe His Gly Gln Ala Leu
    290                 295                 300
Thr Asn Lys Asn Tyr Arg Ile Asp Thr Ile Asn Leu Phe Pro Ala Thr
305                 310                 315                 320
Leu Phe Asp Ala Tyr Met Val Ala Gln Asn Pro Gly Glu Trp Met Leu
                325                 330                 335
Ser Cys Gln Asn Leu Asn His Leu Lys Ala Gly Leu Gln Ala Phe Phe
            340                 345                 350
Gln Val Gln Glu Cys Asn Lys Ser Ser Ser Lys Asp Asn Ile Arg Gly
        355                 360                 365
Lys His Val Arg His Tyr Tyr Ile Ala Ala Glu Glu Ile Ile Trp Asn
    370                 375                 380
Tyr Ala Pro Ser Gly Ile Asp Ile Phe Thr Lys Glu Asn Leu Thr Ala
385                 390                 395                 400
Pro Gly Ser Asp Ser Ala Val Phe Phe Glu Gln Gly Thr Thr Arg Ile
                405                 410                 415
Gly Gly Ser Tyr Lys Lys Leu Val Tyr Arg Glu Tyr Thr Asp Ala Ser
            420                 425                 430
Phe Thr Asn Arg Lys Glu Arg Gly Pro Glu Glu Glu His Leu Gly Ile
        435                 440                 445
Leu Gly Pro Val Ile Trp Ala Glu Val Gly Asp Thr Ile Arg Val Thr
    450                 455                 460
Phe His Asn Lys Gly Ala Tyr Pro Leu Ser Ile Glu Pro Ile Gly Val
465                 470                 475                 480
Arg Phe Asn Lys Asn Asn Glu Gly Thr Tyr Tyr Ser Pro Asn Tyr Asn
                485                 490                 495
Pro Gln Ser Arg Ser Val Pro Pro Ser Ala Ser His Val Ala Pro Thr
            500                 505                 510
Glu Thr Phe Thr Tyr Glu Trp Thr Val Pro Lys Glu Val Gly Pro Thr
        515                 520                 525
Asn Ala Asp Pro Val Cys Leu Ala Lys Met Tyr Tyr Ser Ala Val Asp
    530                 535                 540
Pro Thr Lys Asp Ile Phe Thr Gly Leu Ile Gly Pro Met Lys Ile Cys
545                 550                 555                 560
Lys Lys Gly Ser Leu His Ala Asn Gly Arg Gln Lys Asp Val Asp Lys
                565                 570                 575
Glu Phe Tyr Leu Phe Pro Thr Val Phe Asp Glu Asn Glu Ser Leu Leu
            580                 585                 590
Leu Glu Asp Asn Ile Arg Met Phe Thr Thr Ala Pro Asp Gln Val Asp
        595                 600                 605
Lys Glu Asp Glu Asp Phe Gln Glu Ser Asn Lys Met His Ser Met Asn
    610                 615                 620
Gly Phe Met Tyr Gly Asn Gln Pro Gly Leu Thr Met Cys Lys Gly Asp
625                 630                 635                 640
Ser Val Val Trp Tyr Leu Phe Ser Ala Gly Asn Glu Ala Asp Val His
                645                 650                 655
```

Gly Ile Tyr Phe Ser Gly Asn Thr Tyr Leu Trp Arg Gly Glu Arg Arg
660 665 670
Asp Thr Ala Asn Leu Phe Pro Gln Thr Ser Leu Thr Leu His Met Trp
675 680 685
Pro Asp Thr Glu Gly Thr Phe Asn Val Glu Cys Leu Thr Thr Asp His
690 695 700
Tyr Thr Gly Gly Met Lys Gln Lys Tyr Thr Val Asn Gln Cys Arg Arg
705 710 715 720
Gln Ser Glu Asp Ser Thr Phe Tyr Leu Gly Glu Arg Thr Tyr Tyr Ile
725 730 735
Ala Ala Val Glu Val Glu Trp Asp Tyr Ser Pro Gln Arg Glu Trp Glu
740 745 750
Lys Glu Leu His His Leu Gln Glu Gln Asn Val Ser Asn Ala Phe Leu
755 760 765
Asp Lys Gly Glu Phe Tyr Ile Gly Ser Lys Tyr Lys Lys Val Val Tyr
770 775 780
Arg Gln Tyr Thr Asp Ser Thr Phe Arg Val Pro Val Glu Arg Lys Ala
785 790 795 800
Glu Glu Glu His Leu Gly Ile Leu Gly Pro Gln Leu His Ala Asp Val
805 810 815
Gly Asp Lys Val Lys Ile Ile Phe Lys Asn Met Ala Thr Arg Pro Tyr
820 825 830
Ser Ile His Ala His Gly Val Gln Thr Glu Ser Ser Thr Val Thr Pro
835 840 845
Thr Leu Pro Gly Glu Thr Leu Thr Tyr Val Trp Lys Ile Pro Glu Arg
850 855 860
Ser Gly Ala Gly Thr Glu Asp Ser Ala Cys Ile Pro Trp Ala Tyr Tyr
865 870 875 880
Ser Thr Val Asp Gln Val Lys Asp Leu Tyr Ser Gly Leu Ile Gly Pro
885 890 895
Leu Ile Val Cys Arg Arg Pro Tyr Leu Lys Val Phe Asn Pro Arg Arg
900 905 910
Lys Leu Glu Phe Ala Leu Leu Phe Leu Val Phe Asp Glu Asn Glu Ser
915 920 925
Trp Tyr Leu Asp Asp Asn Ile Lys Thr Tyr Ser Asp His Pro Glu Lys
930 935 940
Val Asn Lys Asp Asp Glu Glu Phe Ile Glu Ser Asn Lys Met His Ala
945 950 955 960
Ile Asn Gly Arg Met Phe Gly Asn Leu Gln Gly Leu Thr Met His Val
965 970 975
Gly Asp Glu Val Asn Trp Tyr Leu Met Gly Met Gly Asn Glu Ile Asp
980 985 990
Leu His Thr Val His Phe His Gly His Ser Phe Gln Tyr Lys His Arg
995 1000 1005
Gly Val Tyr Ser Ser Asp Val Phe Asp Ile Phe Pro Gly Thr Tyr Gln
1010 1015 1020
Thr Leu Glu Met Phe Pro Arg Thr Pro Gly Ile Trp Leu Leu His Cys
1025 1030 1035 1040
His Val Thr Asp His Ile His Ala Gly Met Glu Thr Thr Tyr Thr Val
1045 1050 1055
Leu Gln Asn Glu Asp Thr Lys Ser Gly
1060 1065

<210> 60
<211> 3881
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..(3881)
<223> n = A,T,C or G

<400> 60
agaaagtctt atttataaaa ttgcctgctc ctgattttac ttcatttctt ctcaggctcc 60
aagaaggggga aaaaaatgaa gattttgata cttggtattt ttctgttttt atgtagtacc 120
ccagcctggg cgaaagaaaa gcattattac attggaatta ttgaaacgac ttgggattat 180
gcctctgacc atggggaaaa gaaacttatt tctgttgaca cggaacattc caatatctat 240
cttcaaaatg gcccagatag aattgggaga ctatataaga aggcccttta tcttcagtac 300
acagatgaaa cctttaggac aactatagaa aaaccggtct ggcttgggtt tttaggccct 360
attatcaaag ctgaaactgg agataaagtt tatgtacact taaaaaacct tgcctctagg 420
ccctacacct ttcattcaca tggaataact tactataagg aacatgaggg ggccatctac 480
cctgataaca ccacagattt tcaaagagca gatgacaaag tatatccagg agagcagtat 540
acatacatgt tgcttgccac tgaagaacaa agtcctgggg aaggagatgg caattgtgtg 600
actaggattt accattccca cattgatgct ccaaaagata ttgcctcagg actcatcgga 660
ccttaataa tctgtaaaaa agattctcta gataaagaaa aagaaaaaca tattgaccga 720
gaatttgtgg tgatgtttc tgtggtggat gaaaatttca gctggtacct agaagacaac 780
attaaaacct actgctcaga accagagaaa gttgacaaag acaacgaaga cttccaggag 840
agtaacagaa tgtattctgt gaatggatac actttggaa gtctcccagg actctccatg 900
tgtgctgaag acagagtaaa atggtacctt tttggtatgg gtaatgaagt tgatgtgcac 960
gcagctttct ttcacgggca agcactgact aacaagaact accgtattga cacaatcaac 1020
ctctttcctg ctaccctgtt tgatgcttat atggtggccc agaaccctgg agaatggatg 1080
ctcagctgtc agaatctaaa ccatctgaaa gccggtttgc aagcctttt ccaggttccag 1140
gagtgtaaca agtcttcatc aaaggataat atccgtggga agcatgttag acactactac 1200
attgccgctg aggaaatcat ctggaactat gctccctctg gtatagacat cttcactaaa 1260
gaaaacttaa cagcacctgg aagtgactca gcggtgtttt ttgaacaagg taccacaaga 1320
attggaggct cttataaaaa gcgtggttat cgtgagtaca cagatgcctc cttcacaaat 1380
cgaaaggaga gaggccctga agaagagcat cttggcatcc tgggtcctgt catttgggca 1440
gaggtgggag acaccatcag agtaaccttc cataacaaag gagcatatcc cctcagtatt 1500
gagccgattg gggtgagatt caataagaac aacgagggca catactattc cccaaattac 1560
aaccccaga gcagaagtgt gcctccttca gcctcccatg tggcacccac agaaacattc 1620
acctatgaat ggactgtccc caaagaagta ggacccacta atgcagatcc tgtgtgtcta 1680
gctaagatgt attattctgc tgtggatccc actaaagata tattcactgg gcttattggg 1740
ccaatgaaaa tatgcaagaa aggaagttta catgcaaatg ggagacagaa agatgtagac 1800
aaggaattct attttgtttcc tacagtattt gatgagaatg agagtttact ccatggaagat 1860
aatattagaa tgtttacaac tgcacctgat caggtggata aggaagatga agactttcag 1920
gaatctaata aatgcactc catgaatgga ttcatgtatg ggaactcagcc gggtctcact 1980
atgtgcaaag gagattcggt cgtgtggtac ttattcagcg ccggaaatga ggccgatgta 2040
catggaatat actttcagg aaacacatat ctgtggagag gagaacggag agacacagca 2100
aacctcttcc ctcaaacaag tcttacgctc cacatgtggc ctgacacaga ggggactttt 2160
aatgttgaat gccttacaac tgatcattac acaggcggca tgaagcaaaa atatactgtg 2220
aaccaatgca ggcggcagtc tgaggattcc accttctacc tgggagagag gacatactat 2280
atcgcagcag tggaggtgga atgggattat tccccacaaa gggagtggga aaaggagctg 2340
catcatttac aagagcagaa tgtttcaaat gcatttttag ataagggaga gttttacata 2400
ggctcaaagt acaagaaagt tgtgtatcgg cagtatactg atagcacatt ccgtgttcca 2460
gtggagagaa aagctgaaga agaacatctg ggaattctag gtccacaact tcatgcagat 2520
gttggagaca aagtcaaaat tatctttaaa aacatggcca caaggcccta ctcaatacat 2580
gcccatgggg tacaaacaga gagttctaca gttactccaa cattaccagg tgaaactctc 2640
acttacgtat ggaaaatccc agaaagatct ggagctggaa cagaggattc tgcttgtatt 2700
ccatgggctt attattcaac tgtggatcaa gttaaggacc tctacagtgg attaattggc 2760
cccctgattg tttgtcgaag accttacttg aaagtattca atcccagaag gaagctggaa 2820
tttgcccttc tgtttctagt ttttgatgag aatgaatctt ggtacttaga tgacaacatc 2880
aaaacatact ctgatcaccc cgagaaagta aacaaagatg atgaggaatt catagaaagc 2940
aataaaatgc atgctattaa tggaagaatg tttggaaacc tacaaggcct cacaatgcac 3000
gtgggagatg aagtcaactg gtatctgatg ggaatgggca atgaaataga cttacacact 3060

gtacattttc acggccatag cttccaatac aagcacaggg gagtttatag ttctgatgtc 3120
tttgacattt tccctggaac ataccaaacc ctagaaatgt ttccaagaac acctggaatt 3180
tggttactcc actgccatgt gaccgaccac attcatgctg gaatggaaac cacttacacc 3240
gttctacaaa atgaagcatc ttctgagact cacaggagaa tatggaatgt gatctaccca 3300
atcacagtca gtgtgattat tttattccaa atatctacca aggaatgacc aggagaataa 3360
gatcctccga tgttcgcaat ggtgtggtgt caggaggctg cctcttagac aatctccaga 3420
tgttactgtg atgtgagttt gaaaaagagt tcctgaagta ccacatctgg gagacatgcc 3480
actagctgag cttcccaaaa gtctaccaag agctgaggaa ttgtatcttc atccttgcac 3540
aaagcacctt aaaaacagta aaaggagcct ctatattcca gataaatata gcactgataa 3600
agcgacagct gggcntgaat atcacagcat cacatgggcc acgcaatcaa taatttacca 3660
ttagcattag cgccatggca aaagcaaatt tagacatttt taaaaggaaa cagattctag 3720
gatgtacaaa taaagtttct acagatatag tcttttaatt agtactttta tcttataaac 3780
tgattatgta tgtctgtttt ttcaataaac aaacaaatga aaaaaaaaaa aaaaaaatgg 3840
cggccgcaag cttattancc tttagtgagg gttaatttta a 3881

<210> 61
<211> 1090
<212> PRT
<213> Homo sapiens

<400> 61

```
Met Lys Ile Leu Ile Leu Gly Ile Phe Leu Phe Leu Cys Ser Thr Pro
 1               5                  10                  15
Ala Trp Ala Lys Glu Lys His Tyr Tyr Ile Gly Ile Ile Glu Thr Thr
            20                  25                  30
Trp Asp Tyr Ala Ser Asp His Gly Glu Lys Lys Leu Ile Ser Val Asp
        35                  40                  45
Thr Glu His Ser Asn Ile Tyr Leu Gln Asn Gly Pro Asp Arg Ile Gly
    50                  55                  60
Arg Leu Tyr Lys Lys Ala Leu Tyr Leu Gln Tyr Thr Asp Glu Thr Phe
65                  70                  75                  80
Arg Thr Thr Ile Glu Lys Pro Val Trp Leu Gly Phe Leu Gly Pro Ile
                85                  90                  95
Ile Lys Ala Glu Thr Gly Asp Lys Val Tyr Val His Leu Lys Asn Leu
            100                 105                 110
Ala Ser Arg Pro Tyr Thr Phe His Ser His Gly Ile Thr Tyr Tyr Lys
        115                 120                 125
Glu His Glu Gly Ala Ile Tyr Pro Asp Asn Thr Thr Asp Phe Gln Arg
    130                 135                 140
Ala Asp Asp Lys Val Tyr Pro Gly Glu Gln Tyr Thr Tyr Met Leu Leu
145                 150                 155                 160
Ala Thr Glu Glu Gln Ser Pro Gly Glu Gly Asp Gly Asn Cys Val Thr
                165                 170                 175
Arg Ile Tyr His Ser His Ile Asp Ala Pro Lys Asp Ile Ala Ser Gly
            180                 185                 190
Leu Ile Gly Pro Leu Ile Ile Cys Lys Lys Asp Ser Leu Asp Lys Glu
        195                 200                 205
Lys Glu Lys His Ile Asp Arg Glu Phe Val Val Met Phe Ser Val Val
    210                 215                 220
Asp Glu Asn Phe Ser Trp Tyr Leu Glu Asp Asn Ile Lys Thr Tyr Cys
225                 230                 235                 240
Ser Glu Pro Glu Lys Val Asp Lys Asp Asn Glu Asp Phe Gln Glu Ser
                245                 250                 255
Asn Arg Met Tyr Ser Val Asn Gly Tyr Thr Phe Gly Ser Leu Pro Gly
            260                 265                 270
Leu Ser Met Cys Ala Glu Asp Arg Val Lys Trp Tyr Leu Phe Gly Met
        275                 280                 285
Gly Asn Glu Val Asp Val His Ala Ala Phe Phe His Gly Gln Ala Leu
```

```
    290                 295                 300
Thr Asn Lys Asn Tyr Arg Ile Asp Thr Ile Asn Leu Phe Pro Ala Thr
305                 310                 315                 320
Leu Phe Asp Ala Tyr Met Val Ala Gln Asn Pro Gly Glu Trp Met Leu
                325                 330                 335
Ser Cys Gln Asn Leu Asn His Leu Lys Ala Gly Leu Gln Ala Phe Phe
            340                 345                 350
Gln Val Gln Glu Cys Asn Lys Ser Ser Ser Lys Asp Asn Ile Arg Gly
        355                 360                 365
Lys His Val Arg His Tyr Tyr Ile Ala Ala Glu Glu Ile Ile Trp Asn
    370                 375                 380
Tyr Ala Pro Ser Gly Ile Asp Ile Phe Thr Lys Glu Asn Leu Thr Ala
385                 390                 395                 400
Pro Gly Ser Asp Ser Ala Val Phe Phe Glu Gln Gly Thr Thr Arg Ile
                405                 410                 415
Gly Gly Ser Tyr Lys Lys Leu Val Tyr Arg Glu Tyr Thr Asp Ala Ser
            420                 425                 430
Phe Thr Asn Arg Lys Glu Arg Gly Pro Glu Glu Glu His Leu Gly Ile
        435                 440                 445
Leu Gly Pro Val Ile Trp Ala Glu Val Gly Asp Thr Ile Arg Val Thr
    450                 455                 460
Phe His Asn Lys Gly Ala Tyr Pro Leu Ser Ile Glu Pro Ile Gly Val
465                 470                 475                 480
Arg Phe Asn Lys Asn Asn Glu Gly Thr Tyr Tyr Ser Pro Asn Tyr Asn
                485                 490                 495
Pro Gln Ser Arg Ser Val Pro Pro Ser Ala Ser His Val Ala Pro Thr
            500                 505                 510
Glu Thr Phe Thr Tyr Glu Trp Thr Val Pro Lys Glu Val Gly Pro Thr
        515                 520                 525
Asn Ala Asp Pro Val Cys Leu Ala Lys Met Tyr Tyr Ser Ala Val Asp
    530                 535                 540
Pro Thr Lys Asp Ile Phe Thr Gly Leu Ile Gly Pro Met Lys Ile Cys
545                 550                 555                 560
Lys Lys Gly Ser Leu His Ala Asn Gly Arg Gln Lys Asp Val Asp Lys
                565                 570                 575
Glu Phe Tyr Leu Phe Pro Thr Val Phe Asp Glu Asn Glu Ser Leu Leu
            580                 585                 590
Leu Glu Asp Asn Ile Arg Met Phe Thr Thr Ala Pro Asp Gln Val Asp
        595                 600                 605
Lys Glu Asp Glu Asp Phe Gln Glu Ser Asn Lys Met His Ser Met Asn
    610                 615                 620
Gly Phe Met Tyr Gly Asn Gln Pro Gly Leu Thr Met Cys Lys Gly Asp
625                 630                 635                 640
Ser Val Val Trp Tyr Leu Phe Ser Ala Gly Asn Glu Ala Asp Val His
                645                 650                 655
Gly Ile Tyr Phe Ser Gly Asn Thr Tyr Leu Trp Arg Gly Glu Arg Arg
            660                 665                 670
Asp Thr Ala Asn Leu Phe Pro Gln Thr Ser Leu Thr Leu His Met Trp
        675                 680                 685
Pro Asp Thr Glu Gly Thr Phe Asn Val Glu Cys Leu Thr Thr Asp His
    690                 695                 700
Tyr Thr Gly Gly Met Lys Gln Lys Tyr Thr Val Asn Gln Cys Arg Arg
705                 710                 715                 720
Gln Ser Glu Asp Ser Thr Phe Tyr Leu Gly Glu Arg Thr Tyr Tyr Ile
                725                 730                 735
Ala Ala Val Glu Val Glu Trp Asp Tyr Ser Pro Gln Arg Glu Trp Glu
            740                 745                 750
Lys Glu Leu His His Leu Gln Glu Gln Asn Val Ser Asn Ala Phe Leu
```

```
        755                 760                 765
Asp Lys Gly Glu Phe Tyr Ile Gly Ser Lys Tyr Lys Lys Val Val Tyr
    770                 775                 780
Arg Gln Tyr Thr Asp Ser Thr Phe Arg Val Pro Val Glu Arg Lys Ala
785                 790                 795                 800
Glu Glu Glu His Leu Gly Ile Leu Gly Pro Gln Leu His Ala Asp Val
                805                 810                 815
Gly Asp Lys Val Lys Ile Ile Phe Lys Asn Met Ala Thr Arg Pro Tyr
            820                 825                 830
Ser Ile His Ala His Gly Val Gln Thr Glu Ser Ser Thr Val Thr Pro
        835                 840                 845
Thr Leu Pro Gly Glu Thr Leu Thr Tyr Val Trp Lys Ile Pro Glu Arg
    850                 855                 860
Ser Gly Ala Gly Thr Glu Asp Ser Ala Cys Ile Pro Trp Ala Tyr Tyr
865                 870                 875                 880
Ser Thr Val Asp Gln Val Lys Asp Leu Tyr Ser Gly Leu Ile Gly Pro
                885                 890                 895
Leu Ile Val Cys Arg Arg Pro Tyr Leu Lys Val Phe Asn Pro Arg Arg
            900                 905                 910
Lys Leu Glu Phe Ala Leu Leu Phe Leu Val Phe Asp Glu Asn Glu Ser
        915                 920                 925
Trp Tyr Leu Asp Asp Asn Ile Lys Thr Tyr Ser Asp His Pro Glu Lys
    930                 935                 940
Val Asn Lys Asp Asp Glu Glu Phe Ile Glu Ser Asn Lys Met His Ala
945                 950                 955                 960
Ile Asn Gly Arg Met Phe Gly Asn Leu Gln Gly Leu Thr Met His Val
                965                 970                 975
Gly Asp Glu Val Asn Trp Tyr Leu Met Gly Met Gly Asn Glu Ile Asp
            980                 985                 990
Leu His Thr Val His Phe His Gly His Ser Phe Gln Tyr Lys His Arg
        995                 1000                1005
Gly Val Tyr Ser Ser Asp Val Phe Asp Ile Phe Pro Gly Thr Tyr Gln
    1010                1015                1020
Thr Leu Glu Met Phe Pro Arg Thr Pro Gly Ile Trp Leu Leu His Cys
1025                1030                1035                1040
His Val Thr Asp His Ile His Ala Gly Met Glu Thr Thr Tyr Thr Val
                1045                1050                1055
Leu Gln Asn Glu Ala Ser Ser Glu Thr His Arg Arg Ile Trp Asn Val
            1060                1065                1070
Ile Tyr Pro Ile Thr Val Ser Val Ile Ile Leu Phe Gln Ile Ser Thr
        1075                1080                1085
Lys Glu
    1090
```

<210> 62
<211> 969
<212> DNA
<213> Homo sapiens

<400> 62

```
agctttgggg ttgtccctgg acttgtcttg gttccagaac ctgacgaccc ggcgacggcg 60
acgtctcttt tgactaaaag acagtgtcca gtgctccagc ctaggagtct acgggggaccg 120
cctcccgcgc cgccaccatg cccaacttct ctggcaactg gaaaatcatc cgatcggaaa 180
acttcgagga attgctcaaa gtgctgggag tgaatgtgat gctgaggaag attgctgtgg 240
ctgcagcgtc caagccagca gtggagatca aacaggaggg agacactttc tacatcaaaa 300
cctccaccac cgtgcgcacc acagagatta acttcaaggt tggggaggag tttgaggagc 360
agactgtgga tgggaggccc tgtaagagcc tggtgaaatg ggagagtgag aataaaatgg 420
```

tctgtgagca gaagctcctg aagggagagg gccccaagac ctcgtggacc agagaactga 480
ccaacgatgg ggaactgatc ctgaccatga cggcggatga cgttgtgtgc accagggtct 540
acgtccgaga gtgagtggcc acaggtagaa ccgcggccga agcccaccac tggccatgct 600
caccgccctg cttcactgcc ccctccgtcc caccccctcc ttctaggata gcgctcccct 660
taccccagtc acttctgggg gtcactggga tgcctcttgc agggtcttgc tttctttgac 720
ctcttctctc ctcccctaca ccaacaaaga ggaatggctg caagagccca gatcacccat 780
tccgggttca ctccccgcct ccccaagtca gcagtcctag ccccaaacca gcccagagca 840
gggtctctct aaaggggact tgagggcctg agcaggaaag actggccctc tagcttctac 900
cctttgtccc tgtagcctat acagtttaga atatttattt gttaatttta ttaaaatgct 960
ttaaaaaaa 969

<210> 63
<211> 138
<212> PRT
<213> Homo sapiens

<400> 63
Met Pro Asn Phe Ser Gly Asn Trp Lys Ile Ile Arg Ser Glu Asn Phe
1 5 10 15
Glu Glu Leu Leu Lys Val Leu Gly Val Asn Val Met Leu Arg Lys Ile
20 25 30
Ala Val Ala Ala Ala Ser Lys Pro Ala Val Glu Ile Lys Gln Glu Gly
35 40 45
Asp Thr Phe Tyr Ile Lys Thr Ser Thr Thr Val Arg Thr Thr Glu Ile
50 55 60
Asn Phe Lys Val Gly Glu Glu Phe Glu Glu Gln Thr Val Asp Gly Arg
65 70 75 80
Pro Cys Lys Ser Leu Val Lys Trp Glu Ser Glu Asn Lys Met Val Cys
85 90 95
Glu Gln Lys Leu Leu Lys Gly Glu Gly Pro Lys Thr Ser Trp Thr Arg
100 105 110
Glu Leu Thr Asn Asp Gly Glu Leu Ile Leu Thr Met Thr Ala Asp Asp
115 120 125
Val Val Cys Thr Arg Val Tyr Val Arg Glu
130 135

<210> 64
<211> 927
<212> DNA
<213> Homo sapiens

<400> 64
ggaagtttag gttaactgtc ttaaatttcc aaagctgtaa tcattatttt cattctcaaa 60
gtgatggcct tgtgttttgc tcctctcctc cagggccaga ctgagcccag gttgatttca 120
ggcggacacc aatagactcc acagcagctc caggagccca gacaccggcg gccagaagca 180
aggctaggag ctgctgcagc catgtcggcc ctcagcctcc tcattctggg cctgctcacg 240
gcagtgccac ctgccagctg tcagcaaggc ctggggaacc ttcagccctg gatgcagggc 300
cttatcgcgg tggccgtgtt cctggtcctc gttgcaatcg cctttgcagt caaccactto 360
tggtgccagg aggagccgga gcctgcacac atgatcctga ccgtcggaaa caaggcagat 420
ggagtcctgg tgggaacaga tggaaggtac tcttcgatgg cggccagttt caggtccagt 480
gagcatgaga atgcctatga gaatgtgccc gaggaggaag gcaaggtccg cagcacccag 540
atgtaacctt ctctgtggct ccaaccccaa gactcccagg cacatgggat ggatgtccag 600
tgctaccacc caagcccccc ccttctttgt gtggaatctg caatagtggg ctgactccct 660
ccagccccat gccggcccta cccgcccttg aagtatagcc agccaaggtt ggagctcaga 720
ccgtgtctag gttggggctc ggctgtggcc ctgggtctc ctgctcagct cagaagagcc 780
ttctggagag gacagtcagc tgagcacctc ccatcctgct cacacgtcct tcccctaaac 840
tatggaaatg gccctaattt ctgtgaaata aagacttttt gtatttctgg ggctgaggct 900

cagcaacagc ccctcaggct tccaaaa 927

<210> 65
<211> 114
<212> PRT
<213> Homo sapiens

<400> 65
Met Ser Ala Leu Ser Leu Leu Ile Leu Gly Leu Leu Thr Ala Val Pro
1 5 10 15
Pro Ala Ser Cys Gln Gln Gly Leu Gly Asn Leu Gln Pro Trp Met Gln
20 25 30
Gly Leu Ile Ala Val Ala Val Phe Leu Val Leu Val Ala Ile Ala Phe
35 40 45
Ala Val Asn His Phe Trp Cys Gln Glu Glu Pro Glu Pro Ala His Met
50 55 60
Ile Leu Thr Val Gly Asn Lys Ala Asp Gly Val Leu Val Gly Thr Asp
65 70 75 80
Gly Arg Tyr Ser Ser Met Ala Ala Ser Phe Arg Ser Ser Glu His Glu
85 90 95
Asn Ala Tyr Glu Asn Val Pro Glu Glu Glu Gly Lys Val Arg Ser Thr
100 105 110
Pro Met

<210> 66
<211> 3641
<212> DNA .
<213> Homo sapiens

<400> 66
ctgcactgaa gagggagagc gagagagaga ctggagacgc acagatcccc ccaaggtctc 60
ccaagcctac cgtcccacag attattgtac agagccccaa aaatcgaaac agaggaaacg 120
aacagcagtt gaacatggac gaaggaattc ctcatttgca agagagacag ttactggaac 180
atagagattt tataggactg gactattcct ctttgtatat gtgtaaaccc aaaaggagca 240
tgaaacgaga cgaccaccaag gatacctaca aattaccgca cagattaata gaaaagaaaa 300
gaagagaccg aattaatgaa tgcattgctc agctgaaaga tttactgcct gaacatctga 360
aattgacaac tctgggacat ctggagaaag ctgtagtctt ggaattaact ttgaaacact 420
taaaagcttt aaccgcctta accgagcaac agcatcagga gataattgct ttacagaatg 480
gggagcgatc tctgaaatcg cccattcagt ccgacttgga tgcgttccac tcgggatttc 540
aaacatgcgc caaagaagtc ttgcaatacc tcicccggtt tgagagctgg acacccaggg 600
agccgcggtg tgtccagctg atcaaccact tgcacgccgt ggccacccag ttcttgccca 660
ccccgcagct gttgactcaa caggtccctc tgagcaaagg caccggcgct ccctcggccg 720
ccggtccgc ggccgccccc tgcctggagc gcgcgggcga gaagctggag ccctcgcct 780
actgcgtgcc cgtcatccag cggactcagc ccagcgccga gctcgccgcc gagaacgaca 840
cggacaccga cagcggctac ggcggcgaag ccgaggcccg gccggaccgc gagaaaggca 900
aaggcgcggg ggcgagccgc gtcaccatca agcaggagcc tcccggggag gactcgccgg 960
cgcccaagag gatgaagctg gattcccgcg gcggcggcag cggcggcggc ccgggggggcg 1020
gcgcggcggc ggcggcagcc gcgcttctgg ggcccgaccc tgccgccgcg gccgcgctgc 1080
tgagacccga cgccgccctg ctcagctcgc tggtggcgtt cggcggaggc ggaggcgcgc 1140
ccttcccgca gcccgcggcc gccgcggccc ccttctgcct gcccttctgc ttcctctcgc 1200
cttctgcagc tgccgcctac gtgcagccct tcctggacaa gagcggcctg gagaagtatc 1260
tgtacccggc ggcggctgcc gcccgttcc cgctgctata ccccggcatc cccgccccgg 1320
cggcagccgc ggcagccgcc gccgccgctg ccgccgccgc cgccgcgttc ccctgcctgt 1380
cctggtgttt gtcgcccccc cccgagaagg cgggcgccgc cgccgcgacc ctcctgccgc 1440
acgaggtggc gcccctggg gcgccgcacc cccagcaccc gcacggcagc acccacctgc 1500
ccttcgccgg gcccgcgag ccggggaacc cggagagctc tgctcaggaa gatcctcgc 1560

```
agccaggaaa ggaagctccc tgaatccttg cgtcccgaag gacggaggtt caagcagagt 1620
gagaagttaa aataccctta aggaggttca agcagagtga ggaagttaaaa taccctaag 1680
gtctttaagg gaggaagtgt aatagatgca cgacaggcat aaacaagaac aacaaaacag 1740
gtgttatgtg tacattcgga gttcctgttt tgctcatccc gcaccacccc atcttccaca 1800
cactaacatc cctttcttcc ccccaccagc tgtaaaagat cctatgcgaa agacactggc 1860
tcttttttt aatccccaa ataaattttg ccccctttta ggccatgttc cattatctct 1920
taaaattgga acctaattcg agaggaagta agaagggtct gttctgtggc tgagctaggt 1980
gaacccggg gtaggggaaa gatgttaaca cctttgacgt ctttggagtt gacatggaac 2040
agcaggtagt tgttatgtag agctagttct caaagctgcc ctgcctgttt taggaggcgt 2100
tccacaaaca gattgaggct ctttttagaa ttgaatttac tcttcagtat tttctaatgt 2160
tcagctttct aaaaggcata tatttttcaa agaagtgagg atgcagtttc tcacgttgca 2220
acctattctg aagtggttta aatggtatct cttagtaact tgcactcgtt aaagaaacac 2280
ggagctgggc catcgtcaga actaagtcag ggaaggagat ggatgagaag gccagaatca 2340
ttcctagtac atttgctaac actttattga gaaattgacc atgaattaat ggactcatct 2400
taatttcttc taagtccata tatagataga tatctatctg tacagatttc tatttatcca 2460
tagataggta tctatacata cacatctcaa gtgcatctat tcccactctc attaatccat 2520
catgttccta aatttttgta atcttactgt aaaaaaaagt gcactgaact tcaaaacaaa 2580
acaaaaaaca acaacaacaa aaaacaagtc caaactgata tatcctatat tctgttaaaa 2640
ttcaaaagtg aacgaaagca tttaactggc cagttttgat tgcaaatgct gtaaagatat 2700
agaatgaagt cctgtgaggc cttcctatct ccaagtctat gtatttttctg gagaccaaac 2760
cagataccag ataatcacaa agaaagcttt tttaataagg cttaaaccaa gaccttgtct 2820
agatattttt agtttgttgc caaggtagca ctgtgagaaa tctcacttgg atgttatgta 2880
aggggtgaga cacaacagtc tgactatgag tgaggaaaat atctgggtct tttcgtcagt 2940
ttggtgcatt tgctgctgct gttgctactg tttgcctcaa acgctgtgtt taaacaacgt 3000
taaactctta gcctacaagg tggctcttat gtacatagtt gttaatacat ccaattaatg 3060
atgtctgaca tgctattttt gtagggagaa aatatgtgct aatgatattt tgagttaaaa 3120
tatcttttgg ggaggatttg ctgaaaagtt gcacttttgt tacaatgctt atgcttggta 3180
caagctatg ctgtcttaaa ttattttaaa aaaattaaat actgtctgtg agaaaccagc 3240
tggtttagaa aagtttagta tgtgacgata aactagaaat taccttata ttctagtatt 3300
ttcagcactc cataaattct attacctaaa tattgccaca ctatttgtg atttaaaaat 3360
tcttactaag gaataaaaac tttaatatac gatatgatat tgtctaataa ttaaaaaaga 3420
cataatggat gctcaattag ttttaagata tctataacta tagggataca aatcactaca 3480
gttctcagat ttacaccttt tttttgtcat tggcttgatg tcacacattt ccaatctctt 3540
gcaagcctcc aggctctggc tttgtctacc tgctcgttcc caatgtatct taatgaaaag 3600
tgcaaaagaa aaacctacca attaaaaaaa aaaaaaaaaa a                     3641
```

<210> 67
<211> 482
<212> PRT
<213> Homo sapiens

<400> 67

```
Met Asp Glu Gly Ile Pro His Leu Gln Glu Arg Gln Leu Leu Glu His
 1               5                  10                  15
Arg Asp Phe Ile Gly Leu Asp Tyr Ser Ser Leu Tyr Met Cys Lys Pro
            20                  25                  30
Lys Arg Ser Met Lys Arg Asp Asp Thr Lys Asp Thr Tyr Lys Leu Pro
        35                  40                  45
His Arg Leu Ile Glu Lys Lys Arg Arg Asp Arg Ile Asn Glu Cys Ile
    50                  55                  60
Ala Gln Leu Lys Asp Leu Leu Pro Glu His Leu Lys Leu Thr Thr Leu
65                  70                  75                  80
Gly His Leu Glu Lys Ala Val Val Leu Glu Leu Thr Leu Lys His Leu
                85                  90                  95
Lys Ala Leu Thr Ala Leu Thr Glu Gln Gln His Gln Lys Ile Ile Ala
            100                 105                 110
Leu Gln Asn Gly Glu Arg Ser Leu Lys Ser Pro Ile Gln Ser Asp Leu
        115                 120                 125
```

```
Asp Ala Phe His Ser Gly Phe Gln Thr Cys Ala Lys Glu Val Leu Gln
    130                 135                 140
Tyr Leu Ser Arg Phe Glu Ser Trp Thr Pro Arg Glu Pro Arg Cys Val
145                 150                 155                 160
Gln Leu Ile Asn His Leu His Ala Val Ala Thr Gln Phe Leu Pro Thr
                165                 170                 175
Pro Gln Leu Leu Thr Gln Gln Val Pro Leu Ser Lys Gly Thr Gly Ala
            180                 185                 190
Pro Ser Ala Ala Gly Ser Ala Ala Ala Pro Cys Leu Glu Arg Ala Gly
        195                 200                 205
Gln Lys Leu Glu Pro Leu Ala Tyr Cys Val Pro Val Ile Gln Arg Thr
    210                 215                 220
Gln Pro Ser Ala Glu Leu Ala Ala Glu Asn Asp Thr Asp Thr Asp Ser
225                 230                 235                 240
Gly Tyr Gly Gly Glu Ala Glu Ala Arg Pro Asp Arg Glu Lys Gly Lys
                245                 250                 255
Gly Ala Gly Ala Ser Arg Val Thr Ile Lys Gln Glu Pro Pro Gly Glu
            260                 265                 270
Asp Ser Pro Ala Pro Lys Arg Met Lys Leu Asp Ser Arg Gly Gly Gly
        275                 280                 285
Ser Gly Gly Gly Pro Gly Gly Gly Ala Ala Ala Ala Ala Ala Ala Leu
    290                 295                 300
Leu Gly Pro Asp Pro Ala Ala Ala Ala Ala Leu Leu Arg Pro Asp Ala
305                 310                 315                 320
Ala Leu Leu Ser Ser Leu Val Ala Phe Gly Gly Gly Gly Gly Ala Pro
                325                 330                 335
Phe Pro Gln Pro Ala Ala Ala Ala Ala Pro Phe Cys Leu Pro Phe Cys
            340                 345                 350
Phe Leu Ser Pro Ser Ala Ala Ala Ala Tyr Val Gln Pro Phe Leu Asp
        355                 360                 365
Lys Ser Gly Leu Glu Lys Tyr Leu Tyr Pro Ala Ala Ala Ala Ala Pro
    370                 375                 380
Phe Pro Leu Leu Tyr Pro Gly Ile Pro Ala Pro Ala Ala Ala Ala Ala
385                 390                 395                 400
Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Phe Pro Cys Leu Ser
                405                 410                 415
Ser Val Leu Ser Pro Pro Pro Glu Lys Ala Gly Ala Ala Ala Ala Thr
            420                 425                 430
Leu Leu Pro His Glu Val Ala Pro Leu Gly Ala Pro His Pro Gln His
        435                 440                 445
Pro His Gly Arg Thr His Leu Pro Phe Ala Gly Pro Arg Glu Pro Gly
    450                 455                 460
Asn Pro Glu Ser Ser Ala Gln Glu Asp Pro Ser Gln Pro Gly Lys Glu
465                 470                 475                 480
Ala Pro
```

<210> 68
<211> 3624
<212> DNA
<213> Homo sapiens

<400> 68
tagtctaact cgcggctgtc accgccactg cagcggagcc ggccggccgg gcgctgcggg 60
acgggcgggc ggctgccggc aggaggcgcc gagccgggtg actgccgcgg cgggcacagt 120
ccggggccac agcgccgagc ccgggcggga gtggcccccgc gcaggcaggg agccggccgg 180
cgcactccaa cccggcgggc acctcggggg cgggcgcggg gcgcagcctt ctcgtcccgg 240

cctctgtgac aagcgccccg gagccgggag cccgattgcc gggctcgggg tgggcgcgga 300
cgcaggcact gggctcgtgc ggggccccgg gcgtcgcgat gaacatcgtg gtggagttct 360
tcgtggtcac tttcaaagtg ctctgggcgt tcgtgctggc cgcggcgcgc tggctggtgc 420
ggccCaagga gaagagcgtg gcgggccagg tgtgcctcat cccaggcgcc ggcagcggcc 480
tgggccgcct cttcgcgctg gagttccgcc ggcgtcgggc gctgctggtg ctgtgggaca 540
tcaacacgca aagcaacgag gagacggctg gcatggtgcg ccacatctac cgcgacctgg 600
aggcggccga cgccgctgcg ctgcaagctg ggaatggtga ggaagaaatt ctgccccact 660
gtaacttgca ggtttttacc tacacctgtg acgtggggaa gagggagaac gtctacctga 720
cggctgaaag agtccgcaag gaggttggcg aagtctcagt cctggtcaat aatgctggtg 780
tggtctctgg gcatcacctt ctggaatgtc ctgatgagct cattgagaga accatgatgg 840
tcaattgcca tgcacacttc tggaccacta aggcttttct tcctacgatg ctggagatta 900
atcatggtca tattgtgaca gttgcaagtt ccttgggatt gttcagtact gccggagttg 960
aggattactg tgccagtaaa tttggagttg tgggttttca tgaatccctg agccatgaac 1020
taaaggctgc tgaaaaggat ggaattaaaa caaccttggt ttgccccttat cttgtagaca 1080
ctggcatgtt cagaggctgc cgaatcagga aagaaattga gccttttctg ccacctctga 1140
agcctgatta ctgtgtgaag caggccatga aggccatcct cactgaccag cccatgatct 1200
gcactcccCg cctcatgtac atcgtgacct tcatgaagag catcctacca tttgaagcag 1260
ttgtgtgcat gtatcggttc ctaggagcgg acaagtgtat gtaccccttt attgctcaaa 1320
gaaagcaagc cacaaacaat aatgaagcaa aaaatggaat ctaagaatct ttttgtatgg 1380
aatattactt ctatcagaag atgatcaaga tgtttcagtc cagtgcacat cagcattgct 1440
gacattttat ggattctaaa cttgtgttgt ttctttttta aatcaacttt ttaaaaaaat 1500
aaagtgtaaa ttaaccgact agagtacttg gaaaatgtga tcagtacaag tgaacttagg 1560
ttgttgccaa cagggtcctt ttaggcagaa cccagaaacc agtcaaatct gtagagaagc 1620
agtgtgacat cttcaggtta ccattatttt ttaatgagca ggaagtctag aaatgataac 1680
tagactgtat gtttcatgtg tgtgattttt cagaattccc agagtttact cattcttgtt 1740
attaaactct agccagttga catcttcgca atttcaagga ctgatagtgc tgtatttttct 1800
cacgttttct aagtttccgt tttgcaaggc ctaggtgact ttttcatggt gtttgtatgt 1860
ttagctcttt tgaaaaggaa ttttgaaatc tccatcaact gaagtaaatg atgtctgagt 1920
gttacagtwa aggtgaccaa gtctctttct taaagtcaca atgactaaag tattagttga 1980
attttttttt ttttttttga tggagtctcg ctctgtcacc aggctggagt gcagtagcac 2040
aatcacggct cactgcaatc tctgcctccc rgtttcaagt gattctgctg tctcagcctc 2100
ccaagtagct gggactacag gcatgcgcca ccacgcccag ctaatttttg tatttttagt 2160
agagacgggg tttcaccatg ttggtcagga tggtctccat ctcttgacat tgtgatccac 2220
ctgcctggc ctcccaaagt gctgggatta caggcatgag ccactgcacc cagcctgaa 2280
tttttaattt tatctctgat atacttcatt aagtgtctgg agacctaatt atcctaaaag 2340
atcatacatt ttctacctat gaattttgct gcatacagaa agtgcccttt cctcaggaag 2400
ttgctgtgtt tcatttcttt ggatggactc ttatctagaa tacatagcag ctctgcaaag 2460
aaacagtttt taaaaatggg aacttctaca ttgaaaagtc cccatttttg tgctcaactat 2520
gattagtgag aggaagaaat cttattctat ggcatatgta tggaagggtg taaagattct 2580
tttgaaaggt ttattcacat tgtagaacag caaatgacat ttttacagta ttttttgta 2640
aagcaaacta ttttgtgcct tgaatttggt atatgtgtat tagtgaaaca ttgtaaaggt 2700
gaacttctac ctgtgtatct aaatgtatac catccacttg taaatgacta taaactatta 2760
tgtgattgct ttttttttta gaatgtcttg tttaaatagt ggccaatgtt taaggctgtt 2820
aaaataagcc aactttact aattggggag ttttataaat gactgattaa attttaaagaa 2880
ttaacttaca tgcaattgtg tgattattag ttatcagcag tgttgtaagg aaaattattg 2940
tgttttttt tatgatcatt atcccacttt aggtaaagaa aaatattgga atgagatagt 3000
gttgggaaac agacattaac aacctagggt gcctgcactc aaatagccga tgttactgtc 3060
cctagattag agacttgatt aagggcttgt ttgtaccaaa agtgggggaaa caatgccatg 3120
acctgtgttt tagtttggct gcaccacaga tcaaatctgc actgtgtcta catataggaa 3180
aggtcctggt gtgtgctaat gttcccaatg caggacttga ggaagagctc tgttatatgt 3240
ttccattttct ctttatcaaa gataaccaaa ccttatggcc cttataacaa tggaggcact 3300
ggctgcctct taattttcaa tcatggacct aaagaagtac tctgaagggt ctcaacaatg 3360
ccaggtgggg acagatatac tcagagatta tccaggtctg cctcccagcg agcctggagt 3420
acaccagacc ctcctagaga aatctgttat aatttaacaa cccacttatc caccttaaaa 3480
ctgaggaaag tcgtctttac atctaatttt attcttgtgt gttatatact aaacctattt 3540
ctatttttgt ttgttattgc ccttataagg gtgtccatct ccaagttcaa taaactaatt 3600
catttaaaaa aaaaaaaaa aaaa 3624

<210> 69
<211> 341
<212> PRT
<213> Homo sapiens

<400> 69
Met Asn Ile Val Val Glu Phe Phe Val Val Thr Phe Lys Val Leu Trp
1 5 10 15
Ala Phe Val Leu Ala Ala Ala Arg Trp Leu Val Arg Pro Lys Glu Lys
20 25 30
Ser Val Ala Gly Gln Val Cys Leu Ile Thr Gly Ala Gly Ser Gly Leu
35 40 45
Gly Arg Leu Phe Ala Leu Glu Phe Ala Arg Arg Arg Ala Leu Leu Val
50 55 60
Leu Trp Asp Ile Asn Thr Gln Ser Asn Glu Glu Thr Ala Gly Met Val
65 70 75 80
Arg His Ile Tyr Arg Asp Leu Glu Ala Ala Asp Ala Ala Ala Leu Gln
85 90 95
Ala Gly Asn Gly Glu Glu Glu Ile Leu Pro His Cys Asn Leu Gln Val
100 105 110
Phe Thr Tyr Thr Cys Asp Val Gly Lys Arg Glu Asn Val Tyr Leu Thr
115 120 125
Ala Glu Arg Val Arg Lys Glu Val Gly Glu Val Ser Val Leu Val Asn
130 135 140
Asn Ala Gly Val Val Ser Gly His His Leu Leu Glu Cys Pro Asp Glu
145 150 155 160
Leu Ile Glu Arg Thr Met Met Val Asn Cys His Ala His Phe Trp Thr
165 170 175
Thr Lys Ala Phe Leu Pro Thr Met Leu Glu Ile Asn His Gly His Ile
180 185 190
Val Thr Val Ala Ser Ser Leu Gly Leu Phe Ser Thr Ala Gly Val Glu
195 200 205
Asp Tyr Cys Ala Ser Lys Phe Gly Val Val Gly Phe His Glu Ser Leu
210 215 220
Ser His Glu Leu Lys Ala Ala Glu Lys Asp Gly Ile Lys Thr Thr Leu
225 230 235 240
Val Cys Pro Tyr Leu Val Asp Thr Gly Met Phe Arg Gly Cys Arg Ile
245 250 255
Arg Lys Glu Ile Glu Pro Phe Leu Pro Pro Leu Lys Pro Asp Tyr Cys
260 265 270
Val Lys Gln Ala Met Lys Ala Ile Leu Thr Asp Gln Pro Met Ile Cys
275 280 285
Thr Pro Arg Leu Met Tyr Ile Val Thr Phe Met Lys Ser Ile Leu Pro
290 295 300
Phe Glu Ala Val Val Cys Met Tyr Arg Phe Leu Gly Ala Asp Lys Cys
305 310 315 320
Met Tyr Pro Phe Ile Ala Gln Arg Lys Gln Ala Thr Asn Asn Asn Glu
325 330 335
Ala Lys Asn Gly Ile
340

<210> 70
<211> 1428
<212> DNA
<213> Homo sapiens

<400> 70

```
ggcacgaggc ggagacagag acttcacgac tcccagtctc ctcctcgccg cggccgccgc 60
ctcctcctte tctcctccte ctcttcctcc tcctccctcg ctcccacagc catgtctgct 120
tagaccagag cagccccaca gccaactagg gcagctgccg ccgccacaac agcaaggaca 180
gccgctgccg ccgcccgtga gcgatgacag gagtgtttga cagaagggtc cccagcatcc 240
gatccggcga cttccaagct ccgttccaga cgtccgcagc talgcaccat ccgtctcagg 300
aatcgccaac tttgcccgag tcttcagcta ccgattctga ctactacagc cctacggggg 360
gagccccgca cggctactgc tctcctacct cggcttccta tggcaaagct ctcaaccect 420
accagtatca gtatcacggc gtgaacggct ccgccgggag ctacccagcc aaagcttatg 480
ccgactatag ctacgctagc tcctaccacc agtacggcgg cgcctacaac cgcgtcccaa 540
gcgccaccaa ccagccagag aaagaagtga ccgagcccga ggtgagaatg gtgaatggca 600
aaccaaagaa agttcgtaaa cccaggacta tttattccag ctttcagctg gccgcattac 660
agagaaggtt tcagaagact cagtacctcg ccttgccgga acgcgccgag ctggccgcct 720
cgctgggatt gacacaaaca caggtgaaaa tctggtttca gaacaaaaga tccaagatca 780
agaagatcat gaaaaacggg gagatgcccc cggagcacag tcccagctcc agcgacccaa 840
tggcgtgtaa ctcgccgcag tctccagcgg tgtgggagcc ccagggctcg tcccgctcgc 900
tcagccacca ccctcatgcc caccctccga cctccaacca gtccccagcg tccagctacc 960
tggagaactc tgcatcctgg tacacaagtg cagccagctc aatcaattcc cacctgccgc 1020
cgccggcctc cttacagcac ccgctggcgc tggcctccgg gacactctat tagatgggct 1080
gctctctctt actctctttt ttgggactac tgtgttttgc tgttctagaa aatcataaag 1140
aaaggaattc atatggggaa gttcggaaaa ctgaaaaaga ttcatgtgta aagcttttttt 1200
ttgcatgtaa gttattgcat ttcaaaagac ccccctttt tttacagagg actttttttg 1260
cgcaactgtg gacactttca atggtgcctt gaaatctatg acctcaacct ttcaaaagac 1320
ttttttcaat gttattttag ccatgtaaat aagtgtagat agaggaatta aactgtatat 1380
tctggataaa taaaattatt tcgaccatga aaaaaaaaaa aaaaaaaa 1428
```

<210> 71
<211> 289
<212> PRT
<213> Homo sapiens

<400> 71

```
Met Thr Gly Val Phe Asp Arg Arg Val Pro Ser Ile Arg Ser Gly Asp
 1               5                  10                  15
Phe Gln Ala Pro Phe Gln Thr Ser Ala Ala Met His His Pro Ser Gln
            20                  25                  30
Glu Ser Pro Thr Leu Pro Glu Ser Ser Ala Thr Asp Ser Asp Tyr Tyr
        35                  40                  45
Ser Pro Thr Gly Gly Ala Pro His Gly Tyr Cys Ser Pro Thr Ser Ala
    50                  55                  60
Ser Tyr Gly Lys Ala Leu Asn Pro Tyr Gln Tyr Gln Tyr His Gly Val
65                  70                  75                  80
Asn Gly Ser Ala Gly Ser Tyr Pro Ala Lys Ala Tyr Ala Asp Tyr Ser
                85                  90                  95
Tyr Ala Ser Ser Tyr His Gln Tyr Gly Gly Ala Tyr Asn Arg Val Pro
            100                 105                 110
Ser Ala Thr Asn Gln Pro Glu Lys Glu Val Thr Glu Pro Glu Val Arg
        115                 120                 125
Met Val Asn Gly Lys Pro Lys Lys Val Arg Lys Pro Arg Thr Ile Tyr
    130                 135                 140
Ser Ser Phe Gln Leu Ala Ala Leu Gln Arg Arg Phe Gln Lys Thr Gln
145                 150                 155                 160
Tyr Leu Ala Leu Pro Glu Arg Ala Glu Leu Ala Ala Ser Leu Gly Leu
                165                 170                 175
Thr Gln Thr Gln Val Lys Ile Trp Phe Gln Asn Lys Arg Ser Lys Ile
            180                 185                 190
Lys Lys Ile Met Lys Asn Gly Glu Met Pro Pro Glu His Ser Pro Ser
        195                 200                 205
Ser Ser Asp Pro Met Ala Cys Asn Ser Pro Gln Ser Pro Ala Val Trp
```

```
    210                 215                 220
Glu Pro Gln Gly Ser Ser Arg Ser Leu Ser His His Pro His Ala His
225                 230                 235                 240
Pro Pro Thr Ser Asn Gln Ser Pro Ala Ser Ser Tyr Leu Glu Asn Ser
                245                 250                 255
Ala Ser Trp Tyr Thr Ser Ala Ala Ser Ser Ile Asn Ser His Leu Pro
            260                 265                 270
Pro Pro Gly Ser Leu Gln His Pro Leu Ala Leu Ala Ser Gly Thr Leu
        275                 280                 285
Tyr
```

<210> 72
<211> 2036
<212> DNA
<213> Homo sapiens

<400> 72

```
gacagcttga accattcccc tggccagagt ggattcctca gctatggctc cagcttcagc 60
acctcaccca ctggacagag cccatacacc taccagatgc acggcacaac agggttctat 120
caaggaggaa atggactggg caacgcagcc ggtttcggga gtgtgcacca ggactatcct 180
tcctacccag gcttccccca gagccagtac ccccagtatt acggctcatc ctacaaccct 240
ccctacgtcc cggccagcag catctgccct tcgcccctct ccacgtccac ctacgtcctc 300
caggaggcat ctcacaacgt ccccaaccag agttccgagt cacttgctgg tgaatacaac 360
acacacaatg gaccttccac accagcgaaa gagggagaca cagacaggcc gcaccgggcc 420
tccgacggga agctccgagg ccggtctaag aggagcagtg acccgtcccc ggcaggggac 480
aatgcgattg agcgtgtgtt cgtgtgggac ttggatgaga caatcattat ttttcactcc 540
ttactcacgg ggacatttgc atccagatac gggaaggaca ccacgacgtc cgtgcgcatt 600
ggccttatga tggaagagat gatcttcaac cttgcagata cacatctgtt cttcaatgac 660
ctggaggatt gtgaccagat ccacgttgat gacgtctcat cagatgacaa tggccaagat 720
ttaagcacat acaacttctc cgctgacggc ttccacagtt cggccccagg agccaacctg 780
tgcctgggct ctggcgtgca cggcggcgtg gactggatga ggaagctggc cttccgctac 840
cggcgggtga aggagatgta caatacctac aagaacaacg ttggtgggtt gataggcact 900
cccaaaaggg agacctggct acagctccga gctgagctgg aagctctcac agaccctctg 960
ctgacccact ccctgaaggc actaaacctc atcaactccc ggcccaactg tgtcaatgtg 1020
ctggtcacca ccactcaact aattcctgcc ctggccaaag tcctgctata tggcctgggg 1080
tctgtgtttc ctattgagaa catctacagt gcaaccaaga cagggaagga gagctgcttc 1140
gagaggataa tgcagagatt cggcagaaaa gctgtctacg tggtgatcgg tgatggtgtg 1200
gaagaggagc aaggagcgaa aaagcacaac atgcctttct ggcggatatc ctgccacgca 1260
gacctggagg cactgaggca cgccctggag ctggagtatt tatagcagga tcagcagcat 1320
ctccacctgc catctcaccc tcagacccca tcgccttccc cacctcccca ccgagaactc 1380
cagagaccca gatgttggac accaggaagg ggccccacag ccgagacgac gtgtccagtg 1440
accatctcag aagccgtcca tcagtccaaa tgggggttct gagaaggaaa gtacccaaca 1500
ttggcttcgg agtatttgac tttggggaaa agggctggct cggagtctag actcttctgt 1560
aagactcaca gaacaaaagc aaggaattgc caatttgggg ggtgcctggt gatgaggagg 1620
ggatgggttt gtcttgtctt ctttttaatt tatggactag tctcattact ccagaattat 1680
gctcttgtac ctgtgtggct gggtttctta gtcgttggtt tggtttggtt ttttgaactg 1740
gtatgtaggg tggttcacag ttctaatgta agcactctct tctccaagtt gtgctttgtg 1800
gggacaatca ttctttgaac attagagagg aaggcagttc aagctgttga aaagactatt 1860
gcttatttttt gtttttaaag acctacttga cgtcatgtgg acagtgcacg tgcctacgc 1920
tacatcttgt tttctaggaa gaggggggatg ctgggaagga atgggtgctt tgtgatggat 1980
aaaaggcatt aaataaaacc acgtttacat tttgaaaaaa aaaaaaaaaa aaaaaa     2036
```

<210> 73
<211> 434
<212> PRT
<213> Homo sapiens

<400> 73
Asp Ser Leu Asn His Ser Pro Gly Gln Ser Gly Phe Leu Ser Tyr Gly
1 5 10 15
Ser Ser Phe Ser Thr Ser Pro Thr Gly Gln Ser Pro Tyr Thr Tyr Gln
20 25 30
Met His Gly Thr Thr Gly Phe Tyr Gln Gly Gly Asn Gly Leu Gly Asn
35 40 45
Ala Ala Gly Phe Gly Ser Val His Gln Asp Tyr Pro Ser Tyr Pro Gly
50 55 60
Phe Pro Gln Ser Gln Tyr Pro Gln Tyr Tyr Gly Ser Ser Tyr Asn Pro
65 70 75 80
Pro Tyr Val Pro Ala Ser Ser Ile Cys Pro Ser Pro Leu Ser Thr Ser
85 90 95
Thr Tyr Val Leu Gln Glu Ala Ser His Asn Val Pro Asn Gln Ser Ser
100 105 110
Glu Ser Leu Ala Gly Glu Tyr Asn Thr His Asn Gly Pro Ser Thr Pro
115 120 125
Ala Lys Glu Gly Asp Thr Asp Arg Pro His Arg Ala Ser Asp Gly Lys
130 135 140
Leu Arg Gly Arg Ser Lys Arg Ser Ser Asp Pro Ser Pro Ala Gly Asp
145 150 155 160
Asn Glu Ile Glu Arg Val Phe Val Trp Asp Leu Asp Glu Thr Ile Ile
165 170 175
Ile Phe His Ser Leu Leu Thr Gly Thr Phe Ala Ser Arg Tyr Gly Lys
180 185 190
Asp Thr Thr Thr Ser Val Arg Ile Gly Leu Met Met Glu Glu Met Ile
195 200 205
Phe Asn Leu Ala Asp Thr His Leu Phe Phe Asn Asp Leu Glu Asp Cys
210 215 220
Asp Gln Ile His Val Asp Asp Val Ser Ser Asp Asp Asn Gly Gln Asp
225 230 235 240
Leu Ser Thr Tyr Asn Phe Ser Ala Asp Gly Phe His Ser Ser Ala Pro
245 250 255
Gly Ala Asn Leu Cys Leu Gly Ser Gly Val His Gly Gly Val Asp Trp
260 265 270
Met Arg Lys Leu Ala Phe Arg Tyr Arg Arg Val Lys Glu Met Tyr Asn
275 280 285
Thr Tyr Lys Asn Asn Val Gly Gly Leu Ile Gly Thr Pro Lys Arg Glu
290 295 300
Thr Trp Leu Gln Leu Arg Ala Glu Leu Glu Ala Leu Thr Asp Leu Trp
305 310 315 320
Leu Thr His Ser Leu Lys Ala Leu Asn Leu Ile Asn Ser Arg Pro Asn
325 330 335
Cys Val Asn Val Leu Val Thr Thr Thr Gln Leu Ile Pro Ala Leu Ala
340 345 350
Lys Val Leu Leu Tyr Gly Leu Gly Ser Val Phe Pro Ile Glu Asn Ile
355 360 365
Tyr Ser Ala Thr Lys Thr Gly Lys Glu Ser Cys Phe Glu Arg Ile Met
370 375 380
Gln Arg Phe Gly Arg Lys Ala Val Tyr Val Val Ile Gly Asp Gly Val
385 390 395 400
Glu Glu Glu Gln Gly Ala Lys Lys His Asn Met Pro Phe Trp Arg Ile
405 410 415
Ser Cys His Ala Asp Leu Glu Ala Leu Arg His Ala Leu Glu Leu Glu
420 425 430
Tyr Leu

<210> 74
<211> 1907
<212> DNA
<213> Homo sapiens

<400> 74
cggccagata cctcagcgct acctggcgga actggatttc tctcccgcct gccggcctgc 60
ctgccacagc cggactccgc cactccggta gcctcatggc tgcaacctgt gagattagca 120
acatttttag caactacttc agtgcgatgt acagctcgga ggactccacc ctggcctctg 180
ttccccctgc tgccacctttggggccgatg acttggtact gaccctgagc aaccccccaga 240
tgtcattgga gggtacagag aaggccagct ggttggggga acagccccag ttctggtcga 300
agacgcaggt tctggactgg atcagctacc aagtggagaa gaacaagtac gacgcaagcg 360
ccattgactt ctcacgatgt gacatggatg gcgccaccct ctgcaattgt gcccttgagg 420
agctgcgtct ggtctttggg cctctggggg accaactcca tgcccagctg cgagacctca 480
cttccagctc ttctgatgag ctcagttgga tcattgagct gctggagaag gatggcatgg 540
ccttccagga ggccctagac ccagggccct ttgaccaggg cagccccttt gcccaggagc 600
tgctggacga cggtcagcaa gccagccccct accaccccgg cagctgtggc gcaggagccc 660
cctccctgg cagctctgac gtctccaccg cagggactgg tgcttctcgg agctcccact 720
cctcagactc cggtggaagt gacgtggacc tggatcccac tgatggcaag ctcttcccca 780
gcgatggttt tcgtgactgc aagaaggggg atcccaagca cgggaagcgg aaacgaggcc 840
ggccccgaaa gctgagcaaa gagtactggg actgtctcga gggcaagaag agcaagcacg 900
cgcccagagg cacccacctg tgggagttca tccgggacat cctcatccac ccggagctca 960
acgagggcct catgaagtgg gagaatcggc atgaaggcgt cttcaagttc ctgcgctccg 1020
aggctgtggc ccaactatgg ggccaaaaga aaaagaacag caacatgacc tacgagaagc 1080
tgagccgggc catgaggtac tactacaaac gggagatcct ggaacgggtg gatggccggc 1140
gactcgtcta caagtttggc aaaaactcaa gcggctggaa ggaggaagag gttctccaga 1200
gtcggaactg agggttggaa ctatacccgg gaccaaactc acggaccact cgaggcctgc 1260
aaacctcct gggaggacag gcaggccaga tggcccctcc acttgggcaat gctcccagct 1320
gtgctgtgga gagaagctga tgttttggtg tattgtcagc catcgtcctt ggactcggag 1380
actatggcct cgcctcccca ccctcctctt ggaattacaa gccctggggt ttgaagctga 1440
ctttatagct gcaagtgtat ctccttttat ctggtgcctc ctcaaacccca gtctcagaca 1500
cttaaatgca gacaacacct tcttcctgca gacacttgga ctgagccaag gaggcttggg 1560
aggccctagg gagcaccgtg atggagagga cagagcaggg gctccagcac ttctttctgg 1620
actggcgttc acctccctgc tcagtgcttg ggctccacgg gcaggggtca gagcactccc 1680
taatttatgt gctatataaa tatgtcagat gtacatagag atctattttt tctaaaacat 1740
tcccctcccc actcctctcc cacagagtgc tggactgttc caggccctcc agtgggctga 1800
tgctgggacc cttaggatgg ggctcccagc tcctttctcc tgtgaatgga ggcagagacc 1860
tccaataaag tgccttctgg gctttttcta aaaaaaaaaa aaaaaaa 1907

<210> 75
<211> 371
<212> PRT
<213> Homo sapiens

<400> 75
Met Ala Ala Thr Cys Glu Ile Ser Asn Ile Phe Ser Asn Tyr Phe Ser
1 5 10 15
Ala Met Tyr Ser Ser Glu Asp Ser Thr Leu Ala Ser Val Pro Pro Ala
20 25 30
Ala Thr Phe Gly Ala Asp Asp Leu Val Leu Thr Leu Ser Asn Pro Gln
35 40 45
Met Ser Leu Glu Gly Thr Glu Lys Ala Ser Trp Leu Gly Glu Gln Pro
50 55 60
Gln Phe Trp Ser Lys Thr Gln Val Leu Asp Trp Ile Ser Tyr Gln Val
65 70 75 80
Glu Lys Asn Lys Tyr Asp Ala Ser Ala Ile Asp Phe Ser Arg Cys Asp

```
                85                  90                  95
Met Asp Gly Ala Thr Leu Cys Asn Cys Ala Leu Glu Glu Leu Arg Leu
            100                 105                 110
Val Phe Gly Pro Leu Gly Asp Gln Leu His Ala Gln Leu Arg Asp Leu
        115                 120                 125
Thr Ser Ser Ser Ser Asp Glu Leu Ser Trp Ile Ile Glu Leu Leu Glu
    130                 135                 140
Lys Asp Gly Met Ala Phe Gln Glu Ala Leu Asp Pro Gly Pro Phe Asp
145                 150                 155                 160
Gln Gly Ser Pro Phe Ala Gln Glu Leu Leu Asp Asp Gly Gln Gln Ala
                165                 170                 175
Ser Pro Tyr His Pro Gly Ser Cys Gly Ala Gly Ala Pro Ser Pro Gly
            180                 185                 190
Ser Ser Asp Val Ser Thr Ala Gly Thr Gly Ala Ser Arg Ser Ser His
        195                 200                 205
Ser Ser Asp Ser Gly Gly Ser Asp Val Asp Leu Asp Pro Thr Asp Gly
    210                 215                 220
Lys Leu Phe Pro Ser Asp Gly Phe Arg Asp Cys Lys Lys Gly Asp Pro
225                 230                 235                 240
Lys His Gly Lys Arg Lys Arg Gly Arg Pro Arg Lys Leu Ser Lys Glu
                245                 250                 255
Tyr Trp Asp Cys Leu Glu Gly Lys Lys Ser Lys His Ala Pro Arg Gly
            260                 265                 270
Thr His Leu Trp Glu Phe Ile Arg Asp Ile Leu Ile His Pro Glu Leu
        275                 280                 285
Asn Glu Gly Leu Met Lys Trp Glu Asn Arg His Glu Gly Val Phe Lys
    290                 295                 300
Phe Leu Arg Ser Glu Ala Val Ala Gln Leu Trp Gly Gln Lys Lys Lys
305                 310                 315                 320
Asn Ser Asn Met Thr Tyr Glu Lys Leu Ser Arg Ala Met Arg Tyr Tyr
                325                 330                 335
Tyr Lys Arg Glu Ile Leu Glu Arg Val Asp Gly Arg Arg Leu Val Tyr
            340                 345                 350
Lys Phe Gly Lys Asn Ser Ser Gly Trp Lys Glu Glu Glu Val Leu Gln
        355                 360                 365
Ser Arg Asn
    370
```

<210> 76
<211> 3951
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(3951)
<223> n = A,T,C or G

<400> 76

```
nngacccacg cgtccgggtg cggtctggac acgtctccgg ggtgggtcgt ccggccttcg 60
atcttagacg aattttacaa tgtgaagttc tgcatagatg ccagtcaacc agatgttgga 120
agctggctca agtacattag attcgctggc tgttatgatc agcacaacct tgttgcatgc 180
cagataaatg atcagatatt ctatagagta gttgcagaca ttgcgccggg agaggagctt 240
ctgctgttca tgaagagcga agactatccc catgaaacta tggcgccgga tatccacgaa 300
gaacggcaat atcgctgcga agactgtgac cagctctttg aatctaaggc tgaactagca 360
gatcaccaaa agtttccatg cagtactcct cactcagcat tttcaatggt tgaagaggac 420
tttcagcaaa aactcgaaag cgagaatgat ctccaagaga tacacacgat ccaggagtgt 480
```

```
aaggaatgtg accaagtttt tcctgatttg caaagcctgg agaaacacat gctgtcacat 540
actgaagaga gggaatacaa gtgtgatcag tgtcccaagg catttaactg gaagtccaat 600
ttaattcgcc accagatgtc acatgacagt ggaaagcact atgaatgtga aaactgtgcc 660
aaggttttca cggaccctag caaccttcag cggcacattc gctctcagca tgtcggtgcc 720
cgggcccatg catgcccgga gtgtggcaaa acgtttgcca cttcgtcggg cctcaaacaa 780
cacaagcaca tccacagcag tgtgaagccc tttatctcat tctctcaatc aatgtaccca 840
tttcctgata gagacttgag atcgttacct ttgaaaatgg aaccccaatc accaggtgaa 900
gtaaagaaac tgcagaaggg cagctctgag tcccccttg atctcaccac taagcgaaag 960
gatgagaagc ccttgactcc agtcccctcc aagcctccag tgacacctgc cacaagccaa 1020
gaccagcccc tggatctaag tatgggcagt aggagtagag ccagtgggac aaagctgact 1080
gagcctcgaa aaaaccacgt gtttgggga aaaaaaggaa gcaacgtcga atcaagacct 1140
gcttcagatg gttccttgca gcatgcaaga cccactcctt tctttatgga ccctatttac 1200
agagtagaga aaagaaaact aactgaccca cttgaagctt taaaagagaa atacttgagg 1260
ccttctccag gattcttgtt tcacccacaa atgtcagcta ttgaaaacat ggcagaaaag 1320
ctagagagct tcagtgccct gaaacctgag gccagtgagc tcttacagtc agtgccctct 1380
atgttcaact tcagggcgcc tcccaatgcc ctgccagaga accttctgcg gaagggaaag 1440
gagcgctata cctgcagata ctgtggcaag atttttccaa ggtctgcaaa cctaacacgg 1500
cacttgagaa cccacacagg agagcagcct tacagatgca aatactgtga cagatcattt 1560
agcatatctt ctaacttgca aaggcatgtt cgcaacatcc acaataaaga gaagccattt 1620
aagtgtcact tatgtgatag gtgttttggt caacaaacca atttagacag acacctaaag 1680
aaacatgaga atgggaacat gtccggtaca gcaacatcgt cgcctcattc tgaactggaa 1740
agtacaggtg cgattctgga tgacaaagaa gatgcttact tcacagaaat tcgaaatttc 1800
attgggaaca gcaaccatgg cagccaatct cccaggaatg tggaggagag aatgaatggc 1860
agtcatttta aagatgaaaa ggctttggtg accagtcaaa attcagactt gctggatgat 1920
gaagaagttg aagatgaggt gttgttagat gaggaggatg aagacaatga tattactgga 1980
aaaacaggaa aggaaccagt gacaagtaat ttacatgaag gaaaccctga ggatgactat 2040
gaagaaacca gtgccctgga gatgagttgc aagacatccc cagtgaggta taaagaggaa 2100
gaatataaaa gtggactttc tgctctagat catataaggc acttcacaga tagcctcaaa 2160
atgaggaaaa tggaagataa tcaatattct gaagctgagc tgtcttcttt tagtacttcc 2220
catgtgccag aggaacttaa gcagccgtta cacagaaagt ccaaatcgca ggcatatgct 2280
atgatgctgt cactgtctga caaggagtcc ctccattcta catcccacag ttcttccaac 2340
gtgtggcaca gtatggccag ggctgcggcg gaatccagtg ctatccagtc cataagccac 2400
gtatgacgtt atcaaggttg accagagtgg gaccaagtcc aacagtagca tggctctttc 2460
atataggact atttacaaga ctgctgagca gaatgcctta taaacctgca gggtcactca 2520
tctaaagtct agtgacctta aactgaatga tttaaaaaag aaaagaaaga aaaaagaaac 2580
tatttattct cgatattttg ttttgcacag caaaggcagc tgctgacttc tggaagatca 2640
atcaatgcga cttaaagtga ttcagtgaaa acaaaaaact tggtgggctg aaggcatctt 2700
ccagtttacc ccaccttagg gtatgggtgg gtgagaaggg cagttgagat ggcagcattg 2760
atatgaatga acactccata gaaactgaat tctcttttgt acaagatcac ctgacatgat 2820
tgggaacagt tgcttttaat tacagattta atttttttct tcgttaaagt tttatgtaat 2880
ttaacccttt gaagacagaa gtagttggat gaaatgcaca gtcaattatt atagaaactg 2940
ataacaggga gtacttgttc cccctttgc cttcttaagt acattgttta aaactaggga 3000
aaaagggtat gtgtatattg taaactatgg atgttaacac ttcaaagagg ttaagtcagt 3060
gargtaacct attcatcacc agtaccgctg taccactaat aaattgtttg ccaaatcctt 3120
gtaataacat cttaatttta gacaatcatg tcactgtttt taatgtttat ttttttgtgt 3180
gtgttgcgtg tatcatgtat ttatttgttg gcaaactatt gtttgttgat taaaatagca 3240
ctgttccagt cagccactac tttatgacgt ctgaggcaca ccccttccg aatttcaagg 3300
accaaggtga cccgacctgt gtatgagagt gccaaatggt gtttggcttt tcttaacatt 3360
cctttttgtt tgtttgtttt gtttccttc ttaatgaact aaatacgaat agatgcaact 3420
tagtttttgt aatactgaaa tncgattcaa ttgtataaac gattataatt tctttcatgg 3480
aagcatgatt cttctgatta aaaactgtac tccatatttt atgctggttg tctgcaagct 3540
tgtgcgatgt tatgttcatg ttaatcctat ttgtaaaatg aagtgttccc aaccttatng 3600
ttaaaagaga gaangtaaat aacagactgt attcagttat tttgcccttt attgaggaac 3660
cagatttgtt ttctttttgt ttgtaatctc attttngaaa taatcagcaa gttgaggtac 3720
tttcttcaaa tgctttgtac aatataaact gttatgcctt tcagtgcatt actatgggag 3780
gagcaactaa aaaataaaga cttacaaaaa ggagtatttt taaagaacaa aaacttgagg 3840
gggggcccgg tacccaattc gccctatagt gagtagtatt acaattcact ggccgtcgtt 3900
ttacaacgtc gtgactggga aaaccctggg ttacccaact taatcgtctt n          3951
```

<210> 77
<211> 718
<212> PRT
<213> Homo sapiens

<400> 77
Met Lys Ser Glu Asp Tyr Pro His Glu Thr Met Ala Pro Asp Ile His
1 5 10 15
Glu Glu Arg Gln Tyr Arg Cys Glu Asp Cys Asp Gln Leu Phe Glu Ser
20 25 30
Lys Ala Glu Leu Ala Asp His Gln Lys Phe Pro Cys Ser Thr Pro His
35 40 45
Ser Ala Phe Ser Met Val Glu Glu Asp Phe Gln Gln Lys Leu Glu Ser
50 55 60
Glu Asn Asp Leu Gln Glu Ile His Thr Ile Gln Glu Cys Lys Glu Cys
65 70 75 80
Asp Gln Val Phe Pro Asp Leu Gln Ser Leu Glu Lys His Met Leu Ser
85 90 95
His Thr Glu Glu Arg Glu Tyr Lys Cys Asp Gln Cys Pro Lys Ala Phe
100 105 110
Asn Trp Lys Ser Asn Leu Ile Arg His Gln Met Ser His Asp Ser Gly
115 120 125
Lys His Tyr Glu Cys Glu Asn Cys Ala Lys Val Phe Thr Asp Pro Ser
130 135 140
Asn Leu Gln Arg His Ile Arg Ser Gln His Val Gly Ala Arg Ala His
145 150 155 160
Ala Cys Pro Glu Cys Gly Lys Thr Phe Ala Thr Ser Ser Gly Leu Lys
165 170 175
Gln His Lys His Ile His Ser Ser Val Lys Pro Phe Ile Ser Phe Ser
180 185 190
Gln Ser Met Tyr Pro Phe Pro Asp Arg Asp Leu Arg Ser Leu Pro Leu
195 200 205
Lys Met Glu Pro Gln Ser Pro Gly Glu Val Lys Lys Leu Gln Lys Gly
210 215 220
Ser Ser Glu Ser Pro Phe Asp Leu Thr Thr Lys Arg Lys Asp Glu Lys
225 230 235 240
Pro Leu Thr Pro Val Pro Ser Lys Pro Pro Val Thr Pro Ala Thr Ser
245 250 255
Gln Asp Gln Pro Leu Asp Leu Ser Met Gly Ser Arg Ser Arg Ala Ser
260 265 270
Gly Thr Lys Leu Thr Glu Pro Arg Lys Asn His Val Phe Gly Gly Lys
275 280 285
Lys Gly Ser Asn Val Glu Ser Arg Pro Ala Ser Asp Gly Ser Leu Gln
290 295 300
His Ala Arg Pro Thr Pro Phe Phe Met Asp Pro Ile Tyr Arg Val Glu
305 310 315 320
Lys Arg Lys Leu Thr Asp Pro Leu Glu Ala Leu Lys Glu Lys Tyr Leu
325 330 335
Arg Pro Ser Pro Gly Phe Leu Phe His Pro Gln Met Ser Ala Ile Glu
340 345 350
Asn Met Ala Glu Lys Leu Glu Ser Phe Ser Ala Leu Lys Pro Glu Ala
355 360 365
Ser Glu Leu Leu Gln Ser Val Pro Ser Met Phe Asn Phe Arg Ala Pro
370 375 380
Pro Asn Ala Leu Pro Glu Asn Leu Leu Arg Lys Gly Lys Glu Arg Tyr
385 390 395 400
Thr Cys Arg Tyr Cys Gly Lys Ile Phe Pro Arg Ser Ala Asn Leu Thr

```
                405                 410                 415
Arg His Leu Arg Thr His Thr Gly Glu Gln Pro Tyr Arg Cys Lys Tyr
            420                 425                 430
Cys Asp Arg Ser Phe Ser Ile Ser Ser Asn Leu Gln Arg His Val Arg
        435                 440                 445
Asn Ile His Asn Lys Glu Lys Pro Phe Lys Cys His Leu Cys Asp Arg
    450                 455                 460
Cys Phe Gly Gln Gln Thr Asn Leu Asp Arg His Leu Lys Lys His Glu
465                 470                 475                 480
Asn Gly Asn Met Ser Gly Thr Ala Thr Ser Ser Pro His Ser Glu Leu
                485                 490                 495
Glu Ser Thr Gly Ala Ile Leu Asp Asp Lys Glu Asp Ala Tyr Phe Thr
            500                 505                 510
Glu Ile Arg Asn Phe Ile Gly Asn Ser Asn His Gly Ser Gln Ser Pro
        515                 520                 525
Arg Asn Val Glu Glu Arg Met Asn Gly Ser His Phe Lys Asp Glu Lys
    530                 535                 540
Ala Leu Val Thr Ser Gln Asn Ser Asp Leu Leu Asp Asp Glu Glu Val
545                 550                 555                 560
Glu Asp Glu Val Leu Leu Asp Glu Glu Asp Glu Asp Asn Asp Ile Thr
                565                 570                 575
Gly Lys Thr Gly Lys Glu Pro Val Thr Ser Asn Leu His Glu Gly Asn
            580                 585                 590
Pro Glu Asp Asp Tyr Glu Glu Thr Ser Ala Leu Glu Met Ser Cys Lys
        595                 600                 605
Thr Ser Pro Val Arg Tyr Lys Glu Glu Glu Tyr Lys Ser Gly Leu Ser
    610                 615                 620
Ala Leu Asp His Ile Arg His Phe Thr Asp Ser Leu Lys Met Arg Lys
625                 630                 635                 640
Met Glu Asp Asn Gln Tyr Ser Glu Ala Glu Leu Ser Ser Phe Ser Thr
                645                 650                 655
Ser His Val Pro Glu Glu Leu Lys Gln Pro Leu His Arg Lys Ser Lys
            660                 665                 670
Ser Gln Ala Tyr Ala Met Met Leu Ser Leu Ser Asp Lys Glu Ser Leu
        675                 680                 685
His Ser Thr Ser His Ser Ser Ser Asn Val Trp His Ser Met Ala Arg
    690                 695                 700
Ala Ala Ala Glu Ser Ser Ala Ile Gln Ser Ile Ser His Val
705                 710                 715
```

<210> 78
<211> 4950
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(4950)
<223> n = A,T,C or G

<400> 78
nngacccacg cgtccgggtg cggtctggac acgtctccgg ggtgggtcgt ccggccttcg 60
atcttagacg aattttacaa tgtgaagttc tgcatagatg ccagtcaacc agatgttgga 120
agctggctca agtacattag attcgctggc tgttatgatc agcacaacct tgttgcatgc 180
cagataaatg atcagatatt ctatagagta gttgcagaca ttgcgccggg agaggagctt 240
ctgctgttca tgaagagcga agactatccc catgaaacta tggcgccgga tatccacgaa 300
gaacggcaat atcgctgcga agactgtgac cagctctttg aatctaaggc tgaactagca 360

gatcaccaaa agtttccatg cagtactcct cactcagcat tttcaatggt tgaagaggac 420
tttcagcaaa aactcgaaag cgagaatgat ctccaagaga tacacacgat ccaggagtgt 480
aaggaatgtg accaagtttt tcctgatttg caaagcctgg agaaacacat gctgtcacat 540
actgaagaga gggaatacaa gtgtgatcag tgtcccaagg catttaactg gaagtccaat 600
ttaattcgcc accagatgtc acatgacagt ggaaagcact atgaatgtga aaactgtgcc 660
aaggttttca cggaccctag caaccttcag cggcacattc gctctcagca tgtcggtgcc 720
cgggcccatg catgcccgga gtgtggcaaa acgtttgcca cttcgtcggg cctcaaacaa 780
cacaagcaca tccacagcag tgtgaagccc tttatctgtg aggtctgcca taaatcctat 840
actcagtttt caaacctttg ccgtcataag cgcatgcatg ctgattgcag aacccaaatc 900
aagtgcaaag actgtggaca aatgttcagc actacgtctt ccttaaataa acacaggagg 960
ttttgtgagg gcaagaacca ttttgcggca ggtggatttt ttggccaagg catttcactt 1020
cctggaaccc cagctatgga taaaacgtcc atggttaata tgagtcatgc caacccgggc 1080
cttgctgact attttggcgc caataggcat cctgctggtc ttaccttccc aacagctcct 1140
ggattttctt ttagcttccc tggtctgttt ccttccggct tgtaccacag gcctcctttg 1200
atacctgcta gttctcctgt taaaggacta tcaagtactg aacagacaaa caaaagtcaa 1260
agtcccctca tgacacatcc tcagatactg ccagctacac aggatatttt gaaggcacta 1320
tctaaacacc catctgtagg ggacaataag ccagtggagc tccagcccga gaggtcctct 1380
gaagagaggc cctttgagaa aatcagtgac cagtcagaga gtagtgacct tgatgatgtc 1440
agtacaccaa gtggcagtga cctggaaaca acctcgggct ctgatctgga aagtgacatt 1500
gaaagtgata aagagaaatt taaagaaaat ggtaaaatgt tcaaagacaa agtaagccct 1560
cttcagaatc tggcttcaat aaataataag aaagaataca gcaatcattc catttctca 1620
ccatctttag aggagcagac tgcggtgtca ggagctgtga atgattctat aaaggctatt 1680
gcttctattg ctgaaaaata ctttggttca acaggactgg tggggctgca agacaaaaaa 1740
gttggagctt taccttaccc ttccatgttt cccctcccat tttttccagc attctctcaa 1800
tcaatgtacc catttcctga tagagacttg agatcgttac ctttgaaaat ggaaccccaa 1860
tcaccaggtg aagtaaagaa actgcagaag ggcagctctg agtccccctt tgatctcacc 1920
actaagcgaa aggatgagaa gcccttgact ccagtcccct ccaagcctcc agtgacacct 1980
gccacaagcc aagaccagcc cctggatcta agtatgggca gtaggagtag agccagtggg 2040
acaaagctga ctgagcctcg aaaaaaccac gtgtttgggg gaaaaaaagg aagcaacgtc 2100
gaatcaagac ctgcttcaga tggttccttg cagcatgcaa gacccactcc tttctttatg 2160
gacccctattt acagagtaga gaaaagaaaa ctaactgacc cacttgaagc tttaaaagag 2220
aaatacttga ggccttctcc aggattcttg tttcacccac aattccaact gcctgatcag 2280
agaacttgga tgtcagctat tgaaaacatg gcagaaaagc tagagagctt cagtgccctg 2340
aaacctgagg ccagtgagct cttacagtca gtgccctcta tgttcaactt cagggcgcct 2400
cccaatgccc tgccagagaa ccttctgcgg aagggaaagg agcgctatac ctgcagatac 2460
tgtggcaaga tttttccaag gtctgcaaac ctaacacggc actgagaacc ccacacagga 2520
gagcagcctt acagatgcaa atactgtgac agatcattta gcatatcttc taacttgcaa 2580
aggcatgttc gcaacatcca caataaagag aagccattta agtgtcactt atgtgatagg 2640
tgttttggtc aacaaaccaa tttagacaga cacctaaaga aacatgagaa tgggaacatg 2700
tccggtacag caacatcgtc gcctcattct gaactggaaa gtacaggtgc gattctggat 2760
gacaaagaag atgcttactt cacagaaatt cgaaatttca ttgggaacag caaccatggc 2820
agccaatctc ccaggaatgt ggaggagaga atgaatggca gtcattttaa agatgaaaag 2880
gctttggtga ccagtcaaaa ttcagacttg ctggatgatg aagaagttga agatgaggtg 2940
ttgttagatg aggaggatga agacaatgat attactggaa aaacaggaaa ggaaccagtg 3000
acaagtaatt tacatgaagg aaacccctgag gatgactatg aagaaaccag tgccctggag 3060
atgagttgca agacatcccc agtgaggtat aaagaggaag aatataaaag tggacttcct 3120
gctctagatc atataaggca cttcacagat agcctcaaaa tgaggaaaat ggaagataat 3180
caatattctg aagctgagct gtcttctttt agtacttccc atgtgccaga ggaacttaag 3240
cagccgttac acagaaagtc caaatcgcag gcatatgcta tgatgctgtc actgtctgac 3300
aaggagtccc tccattctac atcccacagt tcttccaacg tgtggcacag tatggccagg 3360
gctgcggcgg aatccagtgc tatccagtcc ataagccacg tatgacgtta tcaaggttga 3420
ccagagtggg accaagtcca acagtagcat ggctctttca tataggacta tttacaagac 3480
tgctgagcag aatgccttat aaacctgcag ggtcactcat ctaaagtcta gtgaccttaa 3540
actgaatgat ttaaaaaaga aaagaaagaa aaaagaaact atttattctc gatattttgt 3600
tttgcacagc aaaggcagct gctgacttct ggaagatcaa tcaatgcgac ttaaagtgat 3660
tcagtgaaaa caaaaaactt ggtgggctga aggcatcttc cagtttaccc caccttaggg 3720
tatgggtggg tgagaagggc agttgagatg gcagccattga tatgaatgaa cactccatag 3780
aaactgaatt ctcttttgta caagatcacc tgacatgatt gggaacagtt gcttttaatt 3840

acagatttaa tttttttctt cgttaaagtt ttatgtaatt taacccttg aagacagaag 3900
tagttggatg aaatgcacag tcaattatta tagaaactga taacagggag tacttgttcc 3960
ccctttgcc ttcttaagta cattgtttaa aactagggaa aaagggtatg tgtatattgt 4020
aaactatgga tgttaacact tcaaagaggt taagtcagtg argtaaccta ttcatcacca 4080
gtaccgctgt accactaata aattgtttgc caaatccttg taataacatc ttaattttag 4140
acaatcatgt cactgttttt aatgtttatt tttttgtgtg tgttgcgtgt atcatgtatt 4200
tatttgttgg caaactattg tttgttgatt aaaatagcac tgttccagtc agccactact 4260
ttatgacgtc tgaggcacac ccctttccga atttcaagga ccaaggtgac ccgacctgtg 4320
tatgagagtg ccaaatggtg tttggctttt cttaacattc cttttttgttt gtttgttttg 4380
ttttccttct taatgaacta aatacgaata gatgcaactt agtttttgta atactgaaat 4440
ncgattcaat tgtataaacg attataattt ctttcatgga agcatgattc ttctgattaa 4500
aaactgtact ccatatttta tgctggttgt ctgcaagctt gtgcgatgtt atgttcatgt 4560
taatcctatt tgtaaaatga agtgttccca accttatngt taaaagagag aangtaaata 4620
acagactgta ttcagttatt ttgcccttta ttgaggaacc agatttgttt tcttttgtt 4680
tgtaatctca ttttngaaat aatcagcaag ttgaggtact ttcttcaaat gctttgtaca 4740
atataaactg ttatgccttt cagtgcatta ctatgggagg agcaactaaa aaataaagac 4800
ttacaaaaag gagtattttt aaagaacaaa aacttgaggg ggggcccggt acccaattcg 4860
ccctatagtg agtagtatta caattcactg gccgtcgttt tacaacgtcg tgactgggaa 4920
aaccctgggt tacccaactt aatcgtcttn 4950

<210> 79
<211> 1051
<212> PRT
<213> Homo sapiens

<400> 79
Met Lys Ser Glu Asp Tyr Pro His Glu Thr Met Ala Pro Asp Ile His
1 5 10 15
Glu Glu Arg Gln Tyr Arg Cys Glu Asp Cys Asp Gln Leu Phe Glu Ser
20 25 30
Lys Ala Glu Leu Ala Asp His Gln Lys Phe Pro Cys Ser Thr Pro His
35 40 45
Ser Ala Phe Ser Met Val Glu Glu Asp Phe Gln Gln Lys Leu Glu Ser
50 55 60
Glu Asn Asp Leu Gln Glu Ile His Thr Ile Gln Glu Cys Lys Glu Cys
65 70 75 80
Asp Gln Val Phe Pro Asp Leu Gln Ser Leu Glu Lys His Met Leu Ser
85 90 95
His Thr Glu Glu Arg Glu Tyr Lys Cys Asp Gln Cys Pro Lys Ala Phe
100 105 110
Asn Trp Lys Ser Asn Leu Ile Arg His Gln Met Ser His Asp Ser Gly
115 120 125
Lys His Tyr Glu Cys Glu Asn Cys Ala Lys Val Phe Thr Asp Pro Ser
130 135 140
Asn Leu Gln Arg His Ile Arg Ser Gln His Val Gly Ala Arg Ala His
145 150 155 160
Ala Cys Pro Glu Cys Gly Lys Thr Phe Ala Thr Ser Ser Gly Leu Lys
165 170 175
Gln His Lys His Ile His Ser Ser Val Lys Pro Phe Ile Cys Glu Val
180 185 190
Cys His Lys Ser Tyr Thr Gln Phe Ser Asn Leu Cys Arg His Lys Arg
195 200 205
Met His Ala Asp Cys Arg Thr Gln Ile Lys Cys Lys Asp Cys Gly Gln
210 215 220
Met Phe Ser Thr Thr Ser Ser Leu Asn Lys His Arg Arg Phe Cys Glu
225 230 235 240
Gly Lys Asn His Phe Ala Ala Gly Gly Phe Phe Gly Gln Gly Ile Ser
245 250 255